Probing Higgs Boson physics in decays to τ leptons with the ATLAS experiment

Brian Le

ORCID.ORG/0000-0002-3917-7079

Doctor of Philosophy

February, 2019

School of Physics
University of Melbourne

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Abstract

The discovery of the Higgs boson in 2012 has opened up a new sector of particle physics to measurements of yet undiscovered couplings. Measurements of the numerous couplings of the Higgs boson - primarily through studies of bosonic decays - have consistently confirmed the Standard Model of Particle Physics. The measurement of fermionic Yukawa couplings, which is a free parameter of the Standard Model, is a critical test of the origin of mass in the Standard Model and serves as a probe for a wide array of new physics models. This thesis presents a measurement, search and feasibility study of couplings of the Higgs boson to the third generation lepton, the τ.

A measurement is presented of the $H \to \tau\tau$ cross-section which is performed using 36.1 fb^{-1} of $\sqrt{s} = 13$ TeV collision data collected at the Large Hadron Collider by the ATLAS detector. A 6.4σ excess over the background-only hypothesis was observed in combination with $\sqrt{s} = 7, 8$ TeV which constitutes a discovery of the final remaining coupling to third generation fermions. The total cross-section was measured as $\sigma_{H\to\tau\tau} = 3.70 \pm 0.58(\text{stat})^{+0.89}_{-0.75}(\text{syst})$ pb, consistent with the expectation from the Standard Model.

Several new physics models predict phenomena which can probed in supressed Higgs boson decays: one such phenomena is lepton flavour violation. A search for lepton flavour violating Higgs boson decays to $e\tau$ and $\mu\tau$ final states, using a new multivariate approach, is also detailed. This search used the same dataset as for the $H \to \tau\tau$ cross-section. No excess is observed and upper limits on the branching ratio for $H \to \mu\tau$ and $H \to e\tau$ of 0.47% and 0.28% are set, respectively. The limit on $H \to e\tau$ is now the new global limit.

Finally a feasibility study for measuring the state of the Higgs boson using decays to two τ leptons is demonstrated. At tree-level, the $H \to \tau\tau$ decays are sensitive to CP admixture couplings to a Higgs boson. One of the promising decay channels, where the τ lepton decays to three charged pions, is shown to be viable thanks to development of a robust neural network approach.

Declaration Page

This is to certify that:

1. this thesis comprises only my original work towards the PhD except where indicated in the Preface;
2. due acknowledgement has been made in the text to all other material used;
3. the thesis is less than 100,000 words in length, exclusive of tables, bibliographies and appendices.

Brian Le

Preface

In recent years, accomplishments in experimental particle physics have been the result of increasingly large international scientific collaborative efforts. This is especially the case for experiments at the Large Hadron Collider, like ATLAS. As a result of these collaborative undertakings, it is incredibly rare that any substantial body of work is done in isolation. The standard procedure is to work in groups, each carrying out a specific purpose within the wider collaboration such as data acquisition, detector monitoring and development, particle reconstruction and event simulation. All these activities represent essential inputs to any scientific achievement in ATLAS. This is the reason why all the members of the collaboration are listed as authors on each ATLAS publication.

The results presented in this thesis have been developed in concert with different working groups and includes some external contributions, each detailed here. The author's original work is contained in chapters 4, 5, 6 and 7 all under the supervision of Prof. E. Barberio and Dr. D. Zanzi.

Chapter 1 contains a brief introduction to the works presented in this thesis. Chapter 2 is an original review of Higgs boson physics designed to motivate the work presented by the author in chapters 5, 6 and 7. Chapter 3 is an original summary of the LHC and ATLAS detectors which are the experimental apparatuses used to collect the data which used in chapters 4, 5 and 6.

Measurement of the τ_{had} energy scale in section 4.4 was performed as part of the ATLAS Tau performance working group in collaboration with Dr. L. Xia. The author contributed to the event selection, background estimation and evaluation of systematic uncertainties with the final signal extraction of the insitu energy scale being performed by Dr. L. Xia. The work presented in this section resulted in [1]. Otherwise this chapter is a summary of τ_{had} reconstruction, identification and calibration which is pertinent to the studies presented in chapters 5, 6 and 7.

Chapter 5 describes the measurement of the $H \to \tau\tau$ cross-section as part of the HLeptons group within ATLAS. The author primarily contributed to the analysis strategy of the semileptonic channel in the event selections, evaluation of systematic uncertainties and signal extraction in collaboration with Dr. A. de Maria and Dr. M. Kassem. In this section only the background estimation, performed by D. Sammel, is included as context for the measurement. The author also developed the signal extraction procedure for the combined measurement with all channels in collaboration with Dr. E. Coniavitis and Dr. M. Hirose. Results from this work have resulted in a conference note [2] and a paper which has been submitted to Physical Review D [3].

Chapter 6 details the author's work in the search for lepton flavour violation (LFV) $H \to e\tau$ and $H \to \mu\tau$ decays as part of the LFV subgroup of the HLeptons group within ATLAS.

The author's contribution to this was the training and evaluation of the BDT method in the non-VBF category of the τ_{had} decay mode, along with the calculation of systematic uncertainties with close consultation with Dr. F. Scutti. Other aspects of this analysis were performed by Dr. X. Chen, Dr. J. Iturbe, Dr. P. Nguyen, Dr. L. Xia and are included to properly contextualise the search results. At the time of submission of this thesis, the search is in the process of internal ATLAS review for publication.

Chapter 7 is an original study of the prospects of using a neural network approach to measure the Higgs boson CP state in decays to τ leptons with three pion final states. The research in this chapter is entirely the author's own work, done independent of ATLAS, in collaboration with Prof. E. Richter-Was, Prof. Z. Was and Dr. D. Zanzi. This work has been published in [4].

Finally chapter 8 summarises the results presented in this thesis.

Acknowledgements

I'm sciencing as fast as I can!
—Prof. Farnsworth, *Futurama*

This PhD has been the longest endeavour of my life and it was not possible alone. There are an endless number of people to thank but I will try to keep it as brief as possible. Firstly I must thank my supervisors Elisabetta Barberio and Daniele Zanzi.

To Elisabetta, you have my most sincere gratitude for taking on a young kid with no experience of experimental physics. Six years ago I was unsure if I had the ability nor the fortitude to pursue experimental physics but you were there to guide and reassure me every step of the way. Your leadership, experience and expertise, especially while working on the $H \to \tau\tau$ analysis, showed me the incredible depth of the field. Thank you in particular for your suggestion to take a beer when I was stressed and not sleeping in Geneva. I now follow that advice as much as possible. I would also like to show my appreciation for the opportunities you've given me by sending me all over the world, it is an experience which I greatly appreciate.

To Daniele, I can only say that it was an absolute pleasure working with you. You always struck the perfect balance of allowing me independence and helping me when I was lost. The incredible patience you've shown whether I was stuck on my code, big picture physics or even collaboration politics: it shows the real character you have. I still remember you telling me you hadn't given up on me, I hope that is still the case after reading my thesis.

I would also like to thank Phil, Rob and Matt who rotated roles on my advisory board panel. Thank you for the patience you've shown in dealing with the beauracratic nightmares.

Of the many postdocs that have come and gone I would particularly like to thank Federico for watching over me for the last couple of years. Despite me making a mess of your framework, you always had time to help fix my code and crash the batch system. Of the many humorous chats we've had, nothing stands out more than your recommendations for good beer. If physics doesn't work out for you, I think we both know where to find you.

During my PhD I have had so many adventures around the world and met so many wonderful people. One opportunity I must thank my supervisors for was the unique opportunity to move to Krakow for a few months to work outside ATLAS. It was an outstanding opportunity to experience another country and to work with some of the best minds in τ physics. I would like to personally thank Zbigniew, Elzbieta, Maria and Andrzej for their hospitality in hosting me and for the wealth of knowledge. Special mention to the secretaries of IFJ-PAN, Beata and Izabela, who helped organise much of my stay in Krakow and who reminded me

when my days off were due! To the friends I made there; Karol, Olga (and little Nadia), Jarek, Szymon, Marzieh, Janka and Jakub (to name a few); thanks for all the good times. I will always remember this wonderful city and will return some day because of you.

Working on a ATLAS has brought me to CERN for much of my PhD and there I have met some amazing individuals. Thanks to Lei, Takashi, Mark, Elias, Minoru, Dirk, Frank, Theo, Chrisitan, Steffen, Michel, Mohammad and Julia for the great conversations (physics or no) during my time working for ATLAS. I hope to work with you all again someday. I would like to especially thank Antonio and Eric who I've spent so many days around CERN with, talking physics or just enjoying a beer. To the CoEPP members who had to share the apartment with me, I miss our random BBQs and I'm sorry for the mess I made in the kitchen (every second day). Marco, Takashi, Millie, Thor, Pere, Noel, Francesca, Francesco, Tom, Carl, Abhishek, Damir and Jason: you were all great housemates to have around and keep things alive in what can often be a dull place.

Back at home, level four has been a second home for me and I've made so many great friends and colleagues. After travelling around the world, I always love coming back and hearing how alive the department feels. From theorists to experiments, Belle to ATLAS you can always find a great conversation (or argument) to make your afternoon (or night) and countless opportunities for a game of pool and intense games of Diplomacy. To Anders, Tristian R., Ash, Peter, Joni, Jo, Eiasha, Tristan B. Justin, John, Tomasz, Leon, Paul, David, KG and many others, it was great spending my mid-20s surrounded by so many vibrant people. A special mention goes to Laurence who was always happy for a endless rant and Giacomo who was always happy for another beer (and free food). I would also like to single out Sean and Lucien for solving so many technical issues directly, or indirectly (sometimes not at all) related to my work. On the whole, I am very grateful to CoEPP for creating an atmosphere that is unmatched and for the funding that allowed the aforementioned travel.

Time always feels like it slips away from you spending endless nights in front of a computer so I would like to *apologise* to my friends and family for never being around. To my parents in particular, I am eternally grateful for the sacrifices you made to make my life easier during this period. Having been away from Melbourne for so long, there were many homesick moments and I was so glad to have you all waiting for me, whenever I would return. I hope that you can be proud of your son for the (few) achievements he's managed.

Contents

1 Introduction **1**

2 The Physics of the Higgs boson **3**
2.1 Standard Model . . . 3
2.1.1 Spontaneous Electroweak Symmetry Breaking and the Higgs boson . 6
2.2 Higgs Boson Production and Decay . . . 9
2.2.1 Production at the LHC . . . 9
2.2.2 Decay Modes . . . 11
2.3 Measurement of Higgs boson Properties . . . 13
2.4 New Physics and the Higgs boson . . . 17
2.4.1 CP Violation in the Higgs sector . . . 17
2.4.2 Charged Lepton Flavour Violation in the Higgs sector . . . 18

3 Experimental Setup **21**
3.1 The Large Hadron Collider . . . 21
3.1.1 Acceleration Process . . . 21
3.1.2 Proton Collisions and Beam Conditions . . . 22
3.1.3 Pileup . . . 24
3.2 The ATLAS Detector . . . 25
3.2.1 Coordinate System of ATLAS . . . 25
3.2.2 Detector description . . . 26
3.2.3 Trigger and Data Acquisition . . . 32
3.2.4 Particle Candidate Reconstruction . . . 33
3.2.5 Monte-Carlo Simulations . . . 37

4 Tau Reconstruction and Calibration **39**
4.1 Reconstruction . . . 41
4.2 Identification . . . 42
4.3 Energy Calibration . . . 45
4.4 Data-driven Corrections . . . 49
4.4.1 Event Selections . . . 49
4.4.2 Background Estimation . . . 51
4.4.3 Efficiency Corrections . . . 54
4.4.4 Tau Energy Scale Extraction . . . 55

5 Measurement of the $H \to \tau\tau$ Cross-section **57**
5.1 Signal and Background . . . 57
5.1.1 Signal Topology . . . 58
5.1.2 Backgrounds . . . 60
5.2 Analysis Strategy . . . 63
5.3 Dataset and MC Simulation . . . 64
5.4 Particle Candidate Selections . . . 66
5.5 Event Selections . . . 67
5.5.1 Signal Regions . . . 67
5.5.2 Background Control Regions . . . 72
5.6 Background Estimation . . . 74
5.6.1 Jets misidentified as τ_{had} Background . . . 74
5.6.2 Top and Diboson Backgrounds . . . 76
5.6.3 $Z \to \tau\tau$ Background . . . 77
5.6.4 Modelling after Background Estimation . . . 80
5.7 Higgs Boson Mass Reconstruction . . . 86
5.7.1 Collinear Mass Approximation . . . 86
5.7.2 Missing Mass Calculator . . . 87
5.8 Systematic Uncertainties . . . 89
5.8.1 Experimental Uncertainties . . . 89
5.8.2 Theoretical Uncertainties . . . 90
5.9 Statistical Hypothesis Testing . . . 93
5.9.1 Test Statistic . . . 93
5.9.2 Hypothesis Testing . . . 94
5.9.3 Statistical Model . . . 95
5.9.4 Fit procedure . . . 97
5.10 Results . . . 99
5.10.1 Semileptonic Channel . . . 99
5.10.2 Combination . . . 101
5.11 Conclusion . . . 104

6 Search for Lepton Flavour Violating Higgs Boson Decays **105**
6.1 Signal and Background . . . 106
6.2 Analysis Strategy . . . 107
6.3 Event Selections . . . 108
6.4 Multivariate Approach . . . 111
6.4.1 Boosted Decision Trees . . . 111
6.4.2 Inputs and Modelling Checks . . . 112
6.4.3 Training Strategy . . . 122
6.5 Background Estimation . . . 126
6.5.1 Backgrounds from Jets Misidentified as τ_{had} . . . 126
6.5.2 Backgrounds from Electron Misidentified as τ_{had} . . . 127
6.6 Fit Model and Systematics . . . 128

6.7 Results . . . 130
6.8 Conclusion and Future Prospects . . . 134

7 Search for CP asymmetry with $H \to \tau\tau$ Decays **135**
7.1 Constructing a CP Sensitive Observable . . . 135
7.1.1 Decay Mode Challenges . . . 137
7.2 Neural Network Approach . . . 141
7.2.1 Inputs and NN Setup . . . 142
7.2.2 Limitations of the Neural Network Approach . . . 144
7.2.3 Detector Resolution Effects . . . 144
7.2.4 Systematic Uncertainty due to τ Modelling . . . 146
7.3 Conclusion and Outlook . . . 152

8 Conclusion **153**

Appendices **169**

A Tau Decay Mode Identification **171**
A.1 Neutral Pion Identification . . . 171
A.2 Decay Mode Classification . . . 172

B LFV MVA performance **175**

List of Figures

2.1 Scalar potential of the Higgs boson [16]. . . . 7
2.2 Feynman diagrams of Higgs boson production at LHC. . . . 9
2.3 Production cross-section of the Higgs boson [18]. . . . 10
2.4 Higgs boson branching fractions [18]. . . . 11
2.5 Summary of Higgs boson mass measurements [22]. . . . 13
2.6 Best fit signal strengths of the Higgs boson couplings [25]. . . . 14
2.7 Summary of measurements of Higgs boson couplings [25]. . . . 15
2.8 Summary of measurements excluding non-scalar spin-CP hypotheses [32]. . . . 16
2.9 Feynman diagrams of Higgs boson mediated LFV decays of τ leptons 19

3.1 Schematic of the LHC apparatus and the associated experiments [77]. . . . 22
3.2 Luminosity delivered by the LHC [78]. . . . 23
3.3 Luminosity as a function of $< \mu >$ [78]. . . . 24
3.4 The definition of angles for the ATLAS coordinate system. . . . 26
3.5 A cross section of the ATLAS detector [9]. . . . 27
3.6 A cross section of the ATLAS inner detector [9]. . . . 29
3.7 A cross section of the calorimeter system of the ATLAS detector [9]. . . . 30
3.8 A cross section of the muon spectrometer of the ATLAS detector [9]. . . . 31
3.9 A schematic of the ATLAS trigger system as of 2015 [81]. . . . 32
3.10 Efficiency curves for the lowest usable single electron HLT triggers in data . 34
3.11 Efficiency curves for the lowest usable single muon HLT triggers in data . . . 35

4.1 Feynman diagrams illustrating the two types of τ lepton decay. . . . 40
4.2 Branching fractions for decay modes of τ leptons [95]. . . . 40
4.3 The response of the identification BDT [97]. . . . 43
4.4 The receiver operating curve for the identification algorithm [97]. . . . 43
4.5 The receiver operating curve for the electron veto [98]. . . . 44
4.6 The calibration curves for the τ energy scale [97]. . . . 45
4.7 The sources of uncertainty for the calibration of the TES for 1p τ_{had} [97]. . . . 46
4.8 The sources of uncertainty for the calibration of the τ energy scale for 3p τ_{had} [97]. . . . 47
4.9 The p_T resolution of τ_{had} candidates with a given p_T for both baseline and BRT approaches [97]. . . . 48
4.10 The m_{vis} distribution in various control regions. . . . 53
4.11 Correction factors for identification and trigger efficiency [1]. . . . 54

4.12 The m_{vis} distribution with TES correction. 56

5.1 Leading order Feynman diagrams which are relevant to this measurement. . 58
5.2 Schematic of a Higgs boson produced via ggH and decaying to a final state of a electron, τ_{had} and neutrinos 59
5.3 Schematic of a Higgs boson produced via VBF and decaying to a final state of a muon, τ_{had} and neutrinos 60
5.4 Examples of Feynman diagrams which contribute to background sources. . . 61
5.5 Expected proportions of signal and background components in both SR and CR of the semileptonic channel. 72
5.6 Example distributions of the m_{MMC} variable in the W+jets and top enriched control regions. 73
5.7 Plots of various key distributions for the validation of the $Z \to \tau\tau$ in the $Z \to ll$ + jets CR corresponding to the semileptonic channel. 79
5.8 Plots of the preselection region in the muon and electron channels. 81
5.9 Plots of the inclusive boosted region in the muon and electron channels. . . . 83
5.10 Plots of the inclusive VBF region in the muon and electron channels. 85
5.11 Distribution of the Higgs boson mass, reconstructed with the MMC algorithm. 88
5.12 A schematic of the fit model. 96
5.13 Post-fit distributions of the m_{MMC} discriminant for the semileptonic channel. 99
5.14 Best fit values for the cross-section for the combined channel fit. 101
5.15 The leading 25 nuisance parameters ranked by the impact on the best-fit cross-section. 102
5.16 Best fit values for the cross-sections of the ggH and VBF production modes. 104

6.1 Topology of a SM Higgs boson decay compared with a LFV Higgs boson decay. 106
6.2 Yields and background composition of the non-VBF categories. 110
6.3 A example of a single decision tree. 111
6.4 Input and output distributions for the BDT in the $\mu\tau_{had}$ non-VBF region. . . 116
6.5 Input and output distributions for the BDT in the $e\tau_{had}$ non-VBF region. . . 119
6.6 The correlations between input variables for the $H \to e\tau$ search. 120
6.7 The correlations between input variables for the $H \to \mu\tau$ search. 121
6.8 A schematic of the BDT training, optimisation and evaluation procedure. . . 123
6.9 Importance of input features for the BDTs. 124
6.10 ROC curves for the test dataset for the trained BDTs. 125
6.11 The top ten nuisance parameters ranked by their impact on μ. 130
6.12 The distribution of the BDT after the fit. 131
6.13 The upper limit of $H \to e\tau$ and $H \to \mu\tau$ branching ratios. 132
6.14 The upper limits on $|Y_{\ell,\tau}|$, $|Y_{\tau,\ell}|$. 133

7.1 A schematic of the calculation of the ϕ^*_{CP} angle. 138
7.2 Distributions of acoplanarity angles of $H \to \tau\tau$ decays in the $\rho - \rho$ decay mode [147]. 139

7.3 Distributions of acoplanarity angles of $H \to \tau\tau$ decays [147]. 140
7.4 A schematic of a single layer NN. 141
7.5 A schematic of the procedure of boosts and rotations needed to make the system of 4-vector inputs uniformly aligned to reduce trivial symmetries. . . 143
7.6 Invariant masses constructed from $\tau^{\mp} \to a_1^{\mp}\nu \to 3\pi^{\mp}\nu$ decays. 147
7.7 A comparison of decay rates as a function of the acoplanarity angles calculated for $H \to \tau\tau \to \rho^{\mp}\nu\ a_1^{\pm}\nu$ using different parameterisations of the decay model. 148
7.8 A comparison of decay rates as a function of the acoplanarity angles calculated for $H \to \tau\tau \to a_1^{\mp}\nu\ a_1^{\pm}\nu$ using different parameterisations of the decay model. 150

A.1 The classifier output for π^0 identification BDT. 172
A.2 The relative or fractional resolution of the reconstructed π^0 candidates. . . . 172
A.3 The reconstruction efficies for each of the τ identification BDTs and ROC curve. 173
A.4 The relative or fractional resolution of the reconstructed τ_{had} candidates in transverse energy (left), η (centre) and ϕ (right) [97]. The resolution with respect to the baseline reconstruction and the particle flow approach. 174

B.1 2D slices of the hyperparameters scanned over to train the BDTs. This scan is done via the Bayesian Optimisation method for the $e\tau_{had}$ channel 176
B.2 AUC scores calculated on the validation dataset in 2D slices of the hyperparameters scanned over to train the BDTs. 176

List of Tables

2.1 The Fermions of the Standard Model. . . . 4
2.2 The Gauge Bosons of the Standard Model. . . . 4
2.3 Current upper limits on LFV couplings to a Higgs boson from various decay modes [65, 70]. . . . 19

3.1 Selected parameters describing the beam conditions delivered by the LHC during data-taking for datasets used here. [79, 80]. . . . 23
3.2 Performance goals of the ATLAS detector [9]. . . . 28

4.1 Event selections used for tag-and-probe measurements. . . . 50
4.2 Monte Carlo generators for various stages of the simulation for the dominant signal and background processes considered. . . . 51
4.3 The leading systematic uncertainties for the measurement of the insitu TES calibration. . . . 56

5.1 Momentum thresholds and isolation for datasets collected with the electron and muon triggers. . . . 64
5.2 Monte Carlo generators used to describe all background and signal processes 65
5.3 Expected number of background events at various stages of event selection in the semileptonic channel. . . . 70
5.4 Summary of selections defining the preselection, inclusive categories and signal regions the semileptonic channel. . . . 71
5.5 The yields calculated by the fit in all the semileptonic SR. . . . 100
5.6 The expected and observed significances of signal for the combined channel as well as the boosted and VBF categories. . . . 100
5.7 Expected and observed impact on the $\sigma_{H\to\tau\tau}$ by various sources of uncertainty. 103

6.1 Summary of selections for preselection and VBF categories for both the $e\tau_{had}$ and $\mu\tau_{had}$ channels. . . . 109
6.2 Expected and observed impact on the $\Delta\mu$ by various sources of uncertainty. . 129
6.3 Event yields in the non-VBF hadronic channel. . . . 131

7.1 AUC scores for NNs trained on various combinations of input features. . . . 145
7.2 AUC score for NNs trained with $a_1 - \rho$ and $a_1 - a_1$ decays of ditau system and then tested on events generated with alternative parameterisations of the decay model. . . . 151

1 Introduction

The decades long search for the final missing piece of the Standard Model (SM) of particle physics [5–7], the Higgs boson, culminated in its discovery at the Large Hadron Collider (LHC) [8] in 2012 by the ATLAS [9] and CMS experiments. The discovery of this particle has verified a cornerstone of the SM; namely electroweak symmetry breaking. Thus far, the Standard Model remains one of the best tested theories in physics. The measurement of the Higgs boson allows for yet unexplored tests of beyond Standard Model (BSM) theories. Many short-falls of the SM can be explained by extensions to the Higgs boson sector, introducing more complex states or new couplings.

The measurement of the Higgs boson's properties is essential in determining whether it is indeed the predicted SM particle. The mass, spin-charge-parity (spin-CP) quantum numbers and couplings to SM particles are all key measurements to the understanding of Higgs boson physics. Overwhelmingly, the measurements made thus far reveal no significant deviation from predictions of the SM. Studies of dibosonic decays of the Higgs boson (WW^*, ZZ^*, $\gamma\gamma$) allowed for very precise measurements of the mass, its bosonic couplings and spin-CP properties. The fermionic decays, while currently not as sensitive as the dibosonic modes, will be vital in measurements of the Yukawa couplings of the Higgs boson; fundamental constants of the SM. Additionally, fermionic decays of the Higgs boson act as a direct probe to BSM physics such as lepton flavour violation (LFV) and charge-parity (CP) violation in the Higgs sector. In particular, the Higgs boson's coupling to tau (τ) leptons opens many opportunities for precision measurements.

This thesis shall focus on probing Higgs boson properties involving couplings to τ leptons. The original work of the author presented in this thesis comprises: the energy calibration of τ leptons, the measurement of the $H \to \tau\tau$ coupling and cross-section, a search for LFV in Higgs boson decays and a feasibility study for a potential measurement the Higgs boson's CP state. With the exception of the feasibility study in chapter 7 and the calibration measurement in chapter 4, these studies utilise 36.1 fb^{-1} of data collected by the ATLAS experiment from proton-proton collisions at the LHC with a centre-of-mass energy of $\sqrt{s} = 13$ TeV. The calibration measurement in chapter 4 utilises 3 fb^{-1} of the $\sqrt{s} = 13$ TeV dataset. A summary of the structure of this thesis is outlined below.

Chapter 2 will describe the underlying theoretical motivations for the analyses to be presented and chapter 3 is a description of the ATLAS detector. Chapter 4 summarises τ lepton reconstruction, identification and calibration. This includes a data-driven measurement of the τ lepton energy scale (TES).

The measurement detailed in chapter 5 encompasses the $H \to \tau\tau$ coupling with a focus on the semileptonic decays of the two τ leptons. The measurement of the Yukawa coupling between the Higgs boson and τ leptons validates the Higgs mechanism's ability to generate fermion masses.

The remaining chapters (6 and 7) will detail a search and a feasibility study for probing BSM models involving Higgs bosons decaying to leptons. These studies borrow much of the analysis structure of the SM $H \to \tau\tau$ measurement either having identical final states or very similar ones. Chapter 6 focuses on constraining potential lepton flavour violating Higgs decays of $H \to \mu\tau$ and $H \to e\tau$ that have a signature which is nearly identical to semileptonic $H \to \tau\tau$ decays. A feasibility study is presented in chapter 7, discussing the potential for the measurement of the Higgs boson's CP properties using a new novel multivariate approach with a focus on fully hadronic final states of the $H \to \tau\tau$ process.

2

The Physics of the Higgs boson

This chapter contains a summary of the SM with a focus on the theoretical motivations underpinning the existence of the Higgs boson as well as the current experimental findings from LHC measurements. This chapter will conclude by motivating the studies of SM and BSM couplings to τ leptons.

2.1 Standard Model

The Standard Model of Particle Physics [5–7] is a quantum field theory which describes much of the current understanding of sub-nuclear scale physics. The SM describes well the fundamental particles and their interactions through the electromagnetic, strong and weak forces. With the exception of the observation of neutrino masses/mixing, most sub-nuclear phenomena are well described by the SM. Many theoretical issues and experimental/cosmological anomalies do exist however: this motivates research into new physics, beyond the SM. This section briefly describes the SM contents leading into the Higgs mechanism.

Fermions

In the SM, the fundamental matter particles are spin 1/2 fermions. The fermions are comprised of two families of particles: leptons and quarks. Each fermion is charged under one or more of the fundamental forces, which dictates what interactions the fermions are affected by. While all particles are charged under the fundamental symmetry of the weak interaction, only quarks are charged under the symmetry of the strong interaction. Additionally electromagnetic interactions are mediated between electrically charged particles. The particle content of fermions in the SM along with their associated quantum numbers is summarised in table 2.1.

Quarks exist in three sets, known as "generations", with particle content of identical quantum numbers. Each generation consists of a pair of quarks containing one "up-type" quark with electric charge +2/3 and one "down-type" quark with electric charge -1/3. Each quark has an associated "flavour" and distinct mass, ranging from MeV scale (up and down quarks) to 175 GeV (top quark), which distinguishes them from particles of the other generations.

Fermions			Quantum Numbers			
			Y	I_3	Q	C
Quarks						
$\binom{u}{d}_L$	$\binom{c}{s}_L$	$\binom{t}{b}_L$	$\frac{1}{3}$	$\binom{+1/2}{-1/2}$	$\binom{+2/3}{-1/3}$	(R,G,B)
u_R	c_R	t_R	$+\frac{4}{3}$	0	$+\frac{2}{3}$	(R,G,B)
d_R	s_R	b_R	$-\frac{2}{3}$	0	$-\frac{1}{3}$	(R,G,B)
Leptons						
$\binom{\nu_e}{e^-}_L$	$\binom{\nu_\mu}{\mu^-}_L$	$\binom{\nu_\tau}{\tau^-}_L$	-1	$\binom{+1/2}{-1/2}$	$\binom{0}{-1}$	0
e^-_R	μ^-_R	τ^-_R	-2	0	-1	0

Table 2.1: The Fermions of the Standard Model. Subscripts indicate whether they are left-handed $SU(2)_L$ doublets (L) or right-handed singlets. The quantum numbers Y, I_3, Q and C are the hypercharge, third component of isospin, electric charge and colour charge respectively.

Gauge Boson	Mass	Electric Charge	Interaction
γ	0	0	Electromagnetic
$W^\pm$	80.399 GeV	± 1	Weak
Z	91.188 GeV	0	Weak
g	0	0	Strong

Table 2.2: The Gauge Bosons of the Standard Model.

As will be noted in the following section, the strong interaction does not allow free particles which are charged. Hence all quarks are found in bound states called "hadrons".

The lepton family can be divided into two classes, neutral neutrinos and charged leptons (more commonly referred to as "leptons"). Leptons also consist of three generations of particles with a charged lepton (electron, muon or tau) and a corresponding neutrino in each generation. The mass scale of the charged leptons range from 511 keV for electrons to 1.7 GeV for tau leptons. Though massless in the SM, neutrinos have been measured to have mass below 1 eV. Henceforth, though, neutrinos will be assumed to be massless.

Gauge Bosons

Particles in the SM interact via gauge bosons which manifest as a property of local gauge symmetry. In a given field theory, imposing local gauge invariance under a lie group G requires the Lagrangian to be invariant under transformations of the type:

$$\psi \rightarrow e^{-i/2\tau\cdot\theta(x)}\psi \tag{2.1}$$

where τ are the generators of the lie group G. The imposition of local gauge invariance under a lie group creates the need for a vector field A_μ in order to correct derivative terms which contain the transformation:

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + i\sum_a \frac{g}{2}\tau^a A^\mu_a, \tag{2.2}$$

where g is the coupling constant of the interaction.

The SM is symmetric under local gauge transformations of the symmetry group:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y, \tag{2.3}$$

where each component of this overall symmetry carries a quantum number: colour, isospin and hypercharge respectively.

The mediation of force between particles, which are charged under these quantum numbers, is via gauge bosons. The gauge bosons of the SM are summarised in table 2.2.

Strong Interactions

Quantum chromodynamics (QCD) is the theory which currently describes the mediation of the colour charge, associated with the $SU(3)_C$ symmetry group, by gluons. A number of interesting phenomena arise from the theory of QCD. The eight generators of this group produce eight coloured gluons and, as $SU(3)$ is a non-abelian group, the strong interaction allows for the self-interaction of gluons.

Two other phenomena also arise from QCD: asymptotic freedom and confinement. The former is a consequence the running of the coupling constant, α_S, which is lower at high energies. The latter describes the effect this has on coloured particles: free quarks do not exist outside of extreme temperatures and must form bound, colourless hadrons. The hadrons are formed from quark-antiquark pairs (mesons) or quark/antiquark triplet states (baryons).

Electroweak Interactions

The electroweak force is a result of the unification of the weak and electromagnetic interactions. The weak force acts under the non-abelian symmetry group $SU(2)$ and the electromagnetic force acts under the abelian $U(1)_{EM}$ symmetry group. The weak force couples to any left-handed chiral fields and the electromagnetic to any electrically charged particle. These two interactions must be unified as the electric charges within the left-handed doublet fields differ, so the interactions can not be merged by considering the direct product of $SU(2) \otimes U(1)_{EM}$.

Electroweak (EW) interactions, as a whole, act under the symmetry group $SU(2)_L \otimes U(1)_Y$. Interactions of chiral fermion fields are principally mediated by a set of massless fields W_i and B, where W_i are associated to the $SU(2)_L$ symmetry group and B is associated with the $U(1)_Y$ group. However, what is observed are massive gauge bosons $W^{\pm}$ and Z mediating charged and neutral currents respectively for the weak interaction as well as a massless photon (γ). In the SM formulation currently described however, such massive bosons would violate local gauge invariance. In a similar sense fermions are also expected to be massless as each fermionic field consists of a left-handed $SU(2)_L$ doublet and a right-handed singlet, breaking gauge invariance.

As will be discussed in the following section, after spontaneous symmetry breaking, W_i and B bosons form mixed eigenstates of observable massive $W^{\pm}$ and Z bosons. This mechanism also provides a means of describing the generation of fermion masses.

2.1.1 Spontaneous Electroweak Symmetry Breaking and the Higgs boson

In the description of the SM thus far, there is no gauge invariant way to introduce mass terms. Thus a mechanism, known as spontaneous symmetry breaking, must be added to break the symmetry while still retaining gauge invariance of the Lagrangian. The breaking of the electroweak symmetry ($SU(2)_L \otimes U(1)_Y \rightarrow U(1)_{EM}$) provides a mechanism [10–15] for introducing masses to all massive particles in the SM with the exception of neutrino masses. The introduction of a scalar potential to the SM Lagrangian:

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \frac{\lambda}{2}(\phi^\dagger \phi)^2 \tag{2.4}$$

with ϕ as a complex $SU(2)_L$ doublet, allows the spontaneous breaking of the $SU(2)_L \otimes U(1)_Y$ symmetry. In cases where the constant μ^2 is negative, this results in a "Mexican hat potential" allowing for a non-trivial vacuum state such as the one illustrated in figure 2.1.

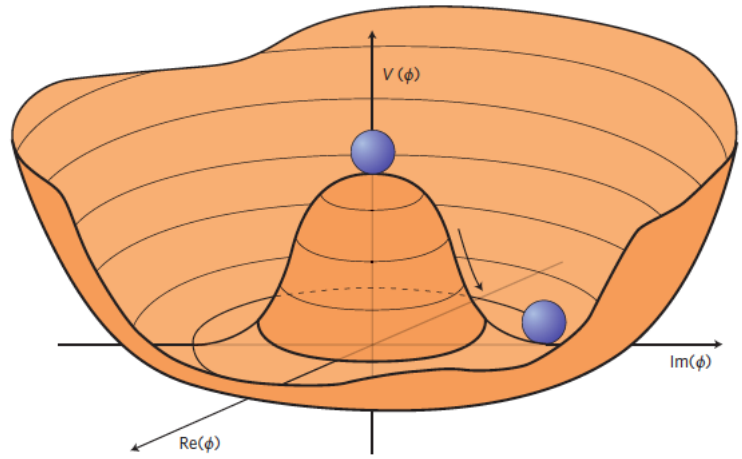


Figure 2.1: An illustration of the scalar potential of the Higgs field with a $\mu^2 < 0$. A degenerate ring of potential vacuum states forms as a minima [16].

The terms in this scalar potential are all invariant under the SM symmetry but the vacuum state of this potential does not obey the same $SU(2)_L \otimes U(1)_Y$ symmetry in the electroweak sector. Without loss of generality, a non-trivial vacuum state (ϕ_0) can be selected

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.5}$$

where v is the free parameter of the vacuum expectation value (VEV). The broken symmetry produces four Goldstone massless scalar bosons according to the Nambu-Goldstone theorem [17]. Under the correct gauge choice, these bosons manifest as longitudinal modes of the electroweak gauge bosons producing the massive gauge bosons ($W^{\pm}$ and Z) and leaving a massless photon:

$$W^{\pm} = \frac{1}{\sqrt{2}}(W_1 \mp iW_2), \tag{2.6}$$

$$Z = \cos\theta_W W_3 - \sin\theta_W B, \tag{2.7}$$

$$\gamma = \sin\theta_W W_3 + \cos\theta_W B, \tag{2.8}$$

where θ_W is the Weinberg mixing angle defined as $\theta_W = \arctan(g/g')$ with g and g' being the weak coupling constants. The masses of the gauge bosons are dictated by three parameters: the VEV v, and the two weak coupling constants g and g'.

$$m_{W^{\pm}} = \frac{1}{2}|g|v, \tag{2.9}$$

$$m_Z = \frac{v}{2}\sqrt{g^2 + g'^2} \tag{2.10}$$

$$m_\gamma = 0. \tag{2.11}$$

The remaining degree of freedom manifests as a singular massive scalar boson, the Higgs boson with a mass of $m_H = \lambda v$.

In addition to the masses of the electroweak sector's gauge bosons, the Higgs boson also provides the masses of the fermions. The direct coupling to fermions is via Yukawa couplings,

$$\mathcal{L}_{\text{Yukawa}} = \sum_f g_f L_f \phi R_f \tag{2.12}$$

where the L and R correspond to the left-handed and right-handed components of each fermion field. The summation is over the various fermionic fields and g_f is the Yukawa coupling for each fermion field. Note that in the SM, there are no right-handed neutrinos and thus neutrinos can not form a mass term in the same way as other fermions.[1]

The masses of other fermions are parameterised with the VEV and g_f.

$$\mathcal{L}_{mass} = \frac{g_f v}{\sqrt{2}} L_f R_f, \tag{2.13}$$

where $g_f v/\sqrt{2}$ is the mass of fermion f. The symmetry breaking mechanism does not explicitly require these terms. The Yukawa couplings must be measured experimentally in order to verify the claim that the Higgs boson generates the masses of fermions through such couplings. To do this, the study of the production and decay modes of the Higgs boson is prudent.

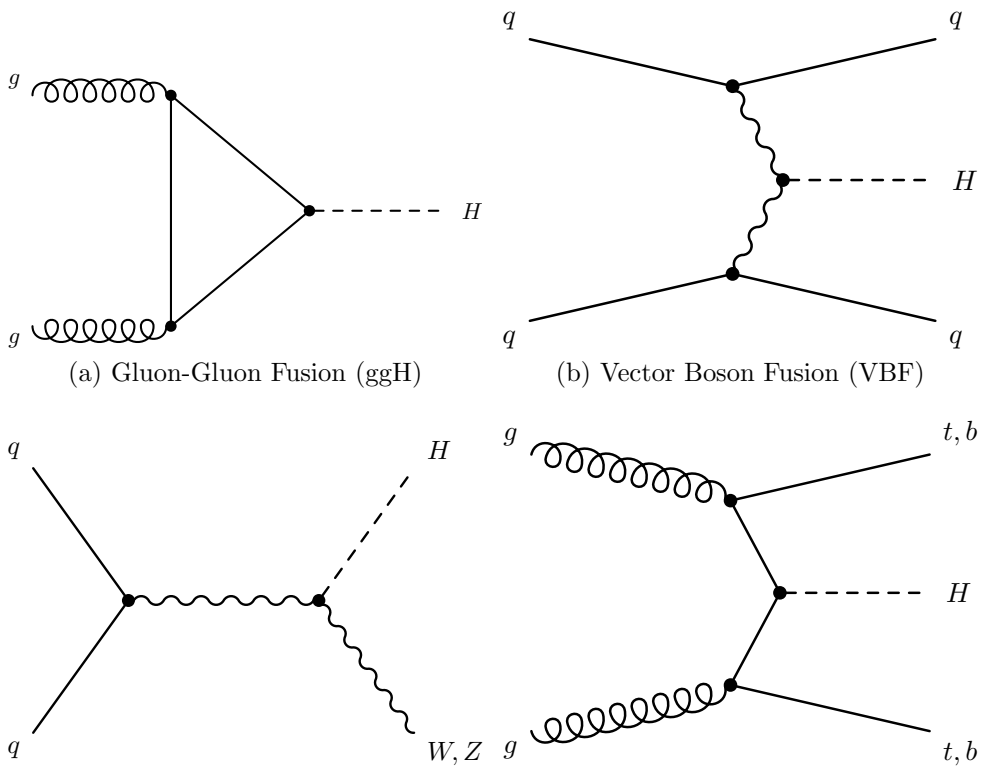
[1] Neutrino masses are confirmed but there has yet been a conclusive theory to explain this phenomena.

2.2 Higgs Boson Production and Decay

Currently the only avenue to study the Higgs boson is through collisions at the LHC, a proton-proton collider operating at a $\sqrt{s} = 13$ TeV. All the Higgs boson couplings have a dependence on mass, which affects both the production modes available at colliders and accessible decay modes. This section will describe the production of the Higgs boson at the LHC and the decay modes which are used to measure the Higgs boson.

2.2.1 Production at the LHC

The production modes are dictated by which particles couple strongly to the Higgs boson, and the rate of such particles in a proton collider. Bosonic (W and Z) couplings and couplings to t-quarks are strongest. In descending order of cross-section, the main production modes of the Higgs boson at the LHC are gluon-gluon fusion (ggH), vector boson fusion (VBF), production in association with a vector boson (VH) and production in association with a $t\bar{t}$ pair (ttH). The leading order Feynman diagrams and cross-sections in p-p collisions are illustrated in figures 2.2 and 2.3 respectively.



(a) Gluon-Gluon Fusion (ggH)

(b) Vector Boson Fusion (VBF)

(c) Associated Production with V (VH)

(d) Associated Production with $t\bar{t}$ (ttH)

Figure 2.2: The leading Feynman diagrams for the major production modes of the Higgs boson at the LHC.

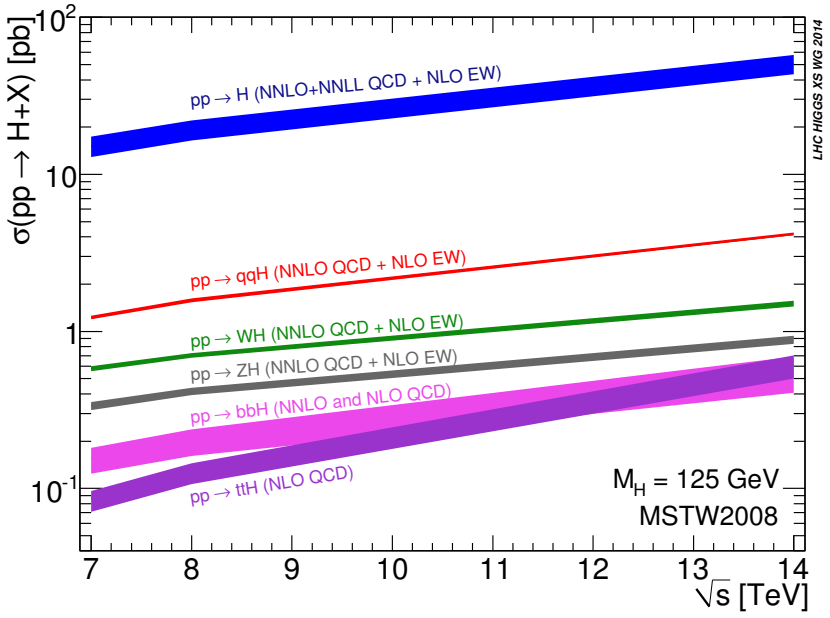


Figure 2.3: Production cross-sections of the Higgs boson at the LHC at different centre of mass energies [18]. The cross-section curves are shown for the major production modes of the LHC as described in the following subsection.

Of these production modes, only the ggH and VBF will be discussed here as they provide the largest sample of Higgs bosons. The VH and ttH production modes are studied in dedicated analyses with the most recent results from ATLAS being respectively: [19] at $\sqrt{s} = 8$ TeV and [20] at $\sqrt{s} = 13$ TeV in combination with other decay modes.

At leading order, the ggH process produces Higgs bosons at rest. However, large corrections from higher order diagrams impact the calculation of the cross-section as well as the topology of events. At next-to-leading order (NLO), the corrections to the leading order cross-section are expected to be greater than 80% [21]. As a result, much of the ggH-produced events are created with an additional jet due to NLO corrections. This produces a boosted topology due to the recoil of the Higgs boson against the additional jet. The next-to-next-to-leading order (NNLO) diagrams impact the cross-section at a reduced rate compared to the NLO corrections (25%) and less still at N^{3}LO corrections. Overall the introduction of higher order QCD corrections is vital to an accurate calculation of the ggH cross-section.

The VBF production mode, conversely, is impacted much less by QCD corrections (5-10%) as it is produced predominantly by electroweak processes. The overall topology of the event consists of two additional jets, radiated by the emission of of the initial-state quarks, produced along with the Higgs boson system. As there is no colour connection between the two outgoing quarks, the radiation of jets is typically collinear with the initial parton direction.

2.2.2 Decay Modes

The Higgs boson has direct couplings to all massive[2] leptons, quarks and gauge bosons. As a result, the Higgs boson has a diverse range of decay modes in which it can be studied. The branching fractions are summarised in figure 2.4 as a function of mass. With a mass of $m_H = 124.97 \pm 0.24$ GeV [22], the leading decay modes of the Higgs boson are $H \to bb$, $H \to \tau\tau$, $H \to W^{\pm}W^{\mp *}$ and $H \to ZZ^*$.

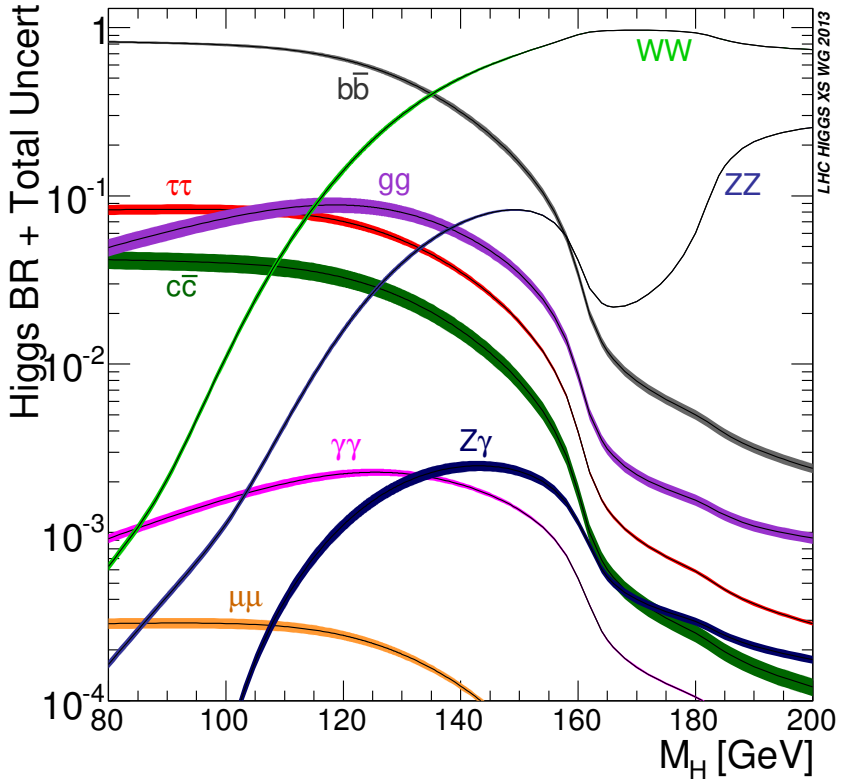


Figure 2.4: Branching fraction for decay modes of the Higgs boson as a function of the mass of a potential Higgs boson [18]. The current best-fit value of the Higgs boson mass is approximately 125 GeV [22].

For the bosonic modes, $H \to W^{\pm}W^{\mp *}$ and $H \to ZZ^*$ decays have the largest branching ratio. The $H \to \gamma\gamma$ channel, despite the low cross-section, is the most experimentally sensitive channel due to a low background. The $H \to ZZ^*$ also has a very distinctive experimental signature with low backgrounds. All three channels are very well established with discoveries in all three channels.

Amongst the fermionic decay modes, the most promising channels are those of $H \to bb$

[2]Recall neutrinos are treated in this thesis as massless

and $H \to \tau\tau$ due to their large branching ratios. The $H \to cc$ is considered a very challenging channel due to the difficulty in identifying hadronic products of a c-quark decay, coupled with the low branching ratio. The current limits on the signal strength set by ATLAS for $H \to cc$ are 110 times the SM expectation [23]. The $H \to \mu\mu$ and $H \to ee$ decays have particularly low branching ratios due to the very low masses of these leptons which reduce the coupling strengths. The latest $H \to \mu\mu$ search performed by ATLAS has set an observed upper limit of 2.8 times the SM expectation [24]. The $H \to ee$ channel is presumably out of reach until the planned High Luminosity LHC upgrade which is expected to increase the available dataset by a factor of 10.

2.3 Measurement of Higgs boson Properties

Measuring the properties of the Higgs boson is one of the goals of the LHC physics programme and of particle physics as a whole. The free parameters of the SM pertaining to the Higgs boson are: the VEV, the coupling constants g and g', the mass of the Higgs boson and the nine Yukawa couplings to fermions. While the mass of the Higgs boson is a completely unmeasured property, the couplings to all massive particles have been be predicted using measurements of the masses of the gauge bosons and fermions. In order to confirm the resonance at ≈125 GeV is indeed the Higgs boson, the measurement of the couplings must be made. In addition to this, the measurement of the Higgs boson's spin-CP state - expected to be a scalar particle - will complete the necessary set of measurements to confirm the SM Higgs boson.

This section describes numerous measurements of the Higgs boson properties at the LHC.

Mass

The mass is best measured in the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels as no information is lost to neutrinos. These channels allow for a complete mass reconstruction from the momenta of final state particles. A summary of the measurements, including the most recent results from $\sqrt{s} = 13$ TeV collisions, are presented in figure 2.5. The current best estimate of the Higgs boson mass is 124.97 ± 0.24 GeV

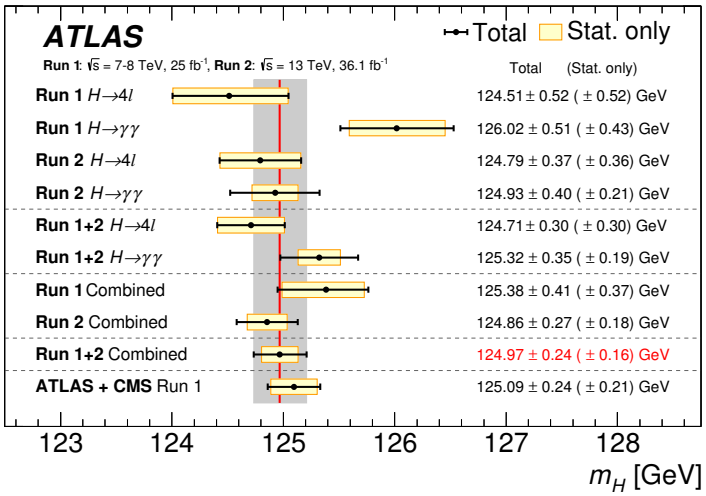


Figure 2.5: Current estimate of the Higgs mass based on results with $\sqrt{s} = 7, 8$ TeV and the latest $\sqrt{s} = 13$ TeVcollisions for the $H \to \gamma\gamma$ and $H \to ZZ^*$ decay modes [22].

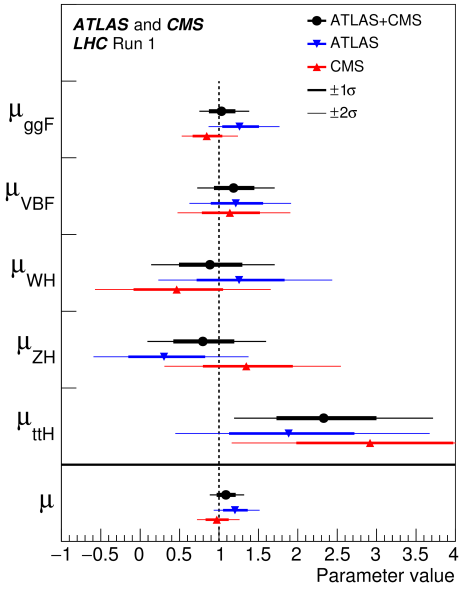
Couplings

Despite the coupling constants being free parameters of the SM, the fact that fermion masses and the VEV have been measured allows predictions to be made regarding the branching fractions of the Higgs boson decay. In the SM, couplings are given by either

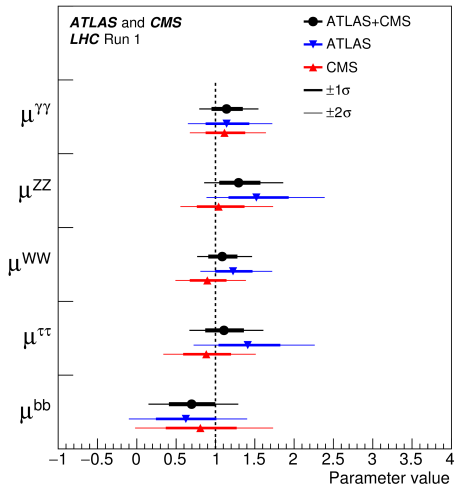
$$g_V \propto 2\frac{m_V^2}{v} \text{ or } g_f = \sqrt{2}vm_f \tag{2.14}$$

where V is either $W^{\pm}$ or Z and f is any massive fermion. Measuring the cross-sections and/or branching fractions across different production and decay modes in a combined fit allows for the extraction of the coupling constants. The measurement is often expressed in terms of the signal strength, μ, which is a ratio of the observed $\sigma \cdot BR$ over the SM expectation.

Overall, combined measurements from ATLAS and CMS experiments from $\sqrt{s} = 7, 8$ TeV collisions [25] indicate no significant deviation from the SM expectation in neither in decay modes nor production modes. Figure 2.6 summarises the best fit signal strength from $\sqrt{s} = 7, 8$ TeV collisions. Figure 2.7 shows the coupling strength with respect to the mass of particles coupling to the Higgs boson. These plots demonstrate good compatibility of measurements with the SM.



(a) Breakdown by production mode



(b) Breakdown by decay mode

Figure 2.6: The best fit result for the signal strengths (μ) from a combination of ATLAS and CMS studies of $\sqrt{s} = 7, 8$ TeV collisions [25].

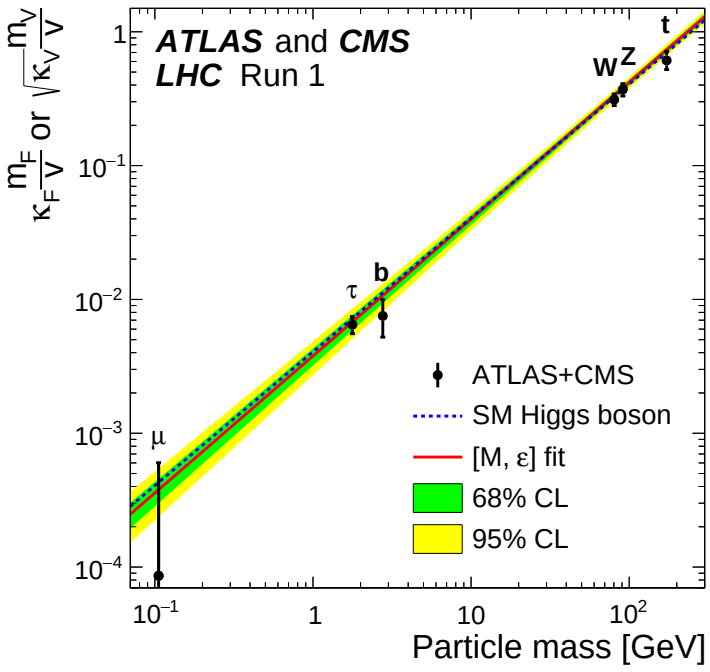


Figure 2.7: The coupling constants as a function of the mass of the particles which couple to the Higgs boson [25]. The κ factor parameterises deviations from the SM coupling strength in a combined fit with all decay modes and production channels, assuming no beyond SM effects [26].

Results derived with $\sqrt{s} = 7, 8$ TeV collisions confirmed the existence of couplings to diboson pairs $\gamma\gamma$, WW^* and ZZ^* [25]. These measurements were not able to confirm couplings to fermions to a sufficient level (5σ) to claim discovery. Given the mass dependence of the Yukawa couplings, couplings to the third generation fermions - the t and b quarks as well as the τ lepton - are expected to be the most sensitive. Following analysis of $\sqrt{s} = 7, 8$ TeV collisions, evidence for a $H \to \tau\tau$ coupling was claimed by ATLAS with a significance of 4.5σ over the no-Higgs boson hypothesis [27]. The aim of chapter 5 measure the total $H \to \tau\tau$ cross-section and observe a 5σ excess in order to claim discovery of the Yukawa coupling to τ leptons. A measurement of the $H \to \tau\tau$ coupling equivalent to the SM prediction will confirm the Higgs mechanism's role in the origin of the τ lepton mass.

As of 2019 measurements have confirmed the existence of couplings to all massive third generation fermions from both ATLAS [3, 20, 28] and CMS [29–31]. Current results, however only measure the fiducial cross-section; measurements of differential cross-sections are still to come. These measurements represent hallmarks of the era of concise Higgs measurements.

Spin-CP

Measurements of the diboson decays of the Higgs boson have firmly excluded all spin states except spin 0 which is predicted by the SM [32]. The measurement of the CP state of the Higgs boson is a more complex matter.

From diboson measurements, it is understood to be a purely scalar particle as predicted by the SM [32]. Figure 2.8 demonstrates that the scalar hypothesis is preferred over any alternatives. The strong limits from diboson decays however neglect that pseudoscalar couplings to bosons can only be added as a dimension 6 operator. This coupling is suppressed by the scale of the new physics compared to tree-level couplings to fermions. For a more comprehensive measurement, fermionic decays are required. These couplings can be added and parameterised in a model independent way. More details on the determination of the Higgs boson CP state with τ lepton decays will be discussed in section 2.4.1 and in chapter 7 as part of the original research presented for this thesis.

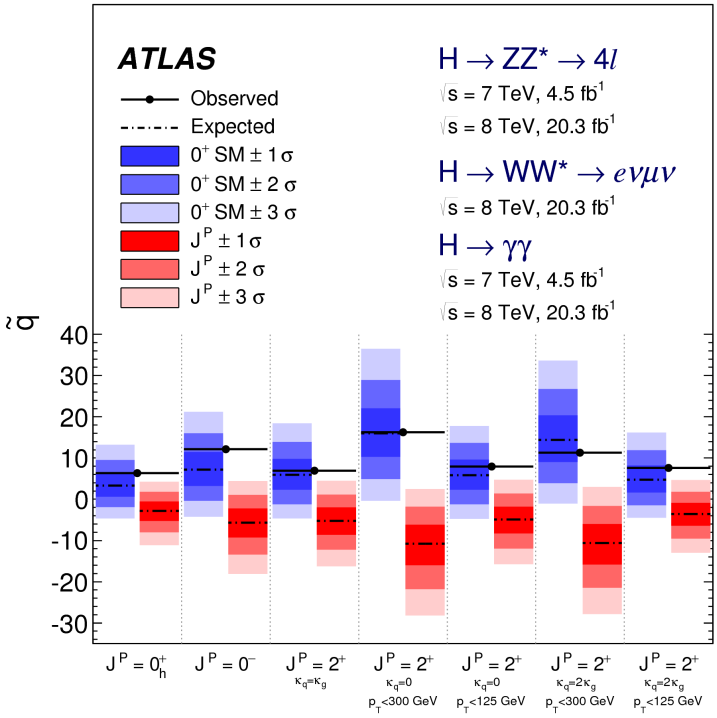


Figure 2.8: Distributions of the test statistic for the SM scalar (blue) and alternative spin-CP (red) hypotheses [32]. The observation favours a scalar hypothesis. Note that the distributions of the SM hypothesis test statistic changes due to different channels being used to exclude certain hypotheses.

2.4 New Physics and the Higgs boson

The SM has several shortcomings: it fails to account for neutrino masses nor does it explain large scale cosmological problems such as baryon asymmetry. Many models have been proposed to alleviate such issues and will often impact the physics of the Higgs boson sector by allowing or requiring new couplings which are not present in the SM.

2.4.1 CP Violation in the Higgs sector

One particularly notable shortcoming of the SM is its inability to explain the baryon asymmetry in the universe. The visible matter consists of baryons which are nearly entirely made of particles rather than anti-particles. The Sakharov conditions [33] are necessary to explain any baryon asymmetry. One condition is the presence of a source of charge-parity (CP) violation. In the SM, small amounts of CP violation are present in the quark sector but this is insufficient to generate the observed asymmetry. One currently unexplored area where an additional source of CP violation can be found is in Higgs boson physics.

Several beyond SM (BSM) scenarios predict the existence of a more complex Higgs boson spectrum. In the SM, there is only a single Higgs doublet however models such as supersymmetry (SUSY) [34] and other two Higgs Doublet (2HDM) models [35–37] predict additional Higgs bosons. One such Higgs boson is predicted to be a pseudoscalar version of the Higgs boson.[3] In these scenarios, it is possible for a pseudoscalar Higgs boson to become degenerate with the scalar SM Higgs boson, producing a mixed CP state. In this case, a CP violating coupling may produce a source of CP violation.

While upper limits have been set on anomalous couplings between a pseudoscalar Higgs boson and two gauge bosons in effective field theories, these studies only concern the diabosonic decay modes which are less sensitive to direct couplings to a pseudoscalar Higgs boson [38–47]. The fermionic decay modes, in which the Yukawa coupling allows direct couplings to pseudoscalar Higgs bosons, are an ideal channel to measure any potential mixings.

Much of the literature proposes the measurement of the CP phase of the Higgs boson in the $\tau\tau$ final state [48–52], particularly in τ decays to one or two pions. CP information is encoded in transverse spin components of the τ leptons which can be measured in angular distributions of decay products. The opportunity of using fermionic decays to perform a measurement of the CP state are now becoming viable, thanks to developments in τ decay reconstruction.[4] Measurement using these decays will ultimately give a direct probe to the CP state of the Higgs boson. Currently however, attempts of measurements are hampered by a lack of an adequately sized sample.

[3] Note a pseudoscalar is a spin 0 particle which changes sign under the parity operator.
[4] See appendix A for a summary.

To bolster the available data, a method has been developed to include complex three pion decays using a multivariate technique. This research is presented in chapter 7.

2.4.2 Charged Lepton Flavour Violation in the Higgs sector

In the SM, each fermion has an associated "flavour" to it, which indicates what type of particle it is. Theoretically, there is some redundancy in having separate generations of particles with no mixing between. It is known that flavour symmetries are often broken within families of fermions. The discovery of neutrino mixing and suppressed hadronic decays, where quarks couple across generations, have been known for decades. This indicates flavour mixing in both the neutrino and quark sectors. The charged leptons are conversely the only remaining unbroken flavour symmetry in the SM: a curious feature. Many extensions to the SM including 2HDM [53, 54], SUSY [55–59], composite Higgs [60, 61] and Randall-Sundrum models [62]. While motivations for these models are quite sparse, from solving issues such as the hierarchy problem to explaining neutrino masses, these can generate lepton flavour violation in Higgs boson decays.

Searches for LFV decays come primarily through searching for $\ell \to \ell'\gamma$ decays as well as $\ell \to 3\ell'$ decays, where ℓ is a charged lepton and $\ell \neq \ell'$. Any excesses in measurements of these decays, which are highly suppressed in the SM, would indicate physics beyond the SM. To date, precision measurements have been performed to find the highly suppressed $\mu \to e\gamma$ transition using highly energetic muon beams. The latest measurement from the MEG collaboration constrains the branching ratio $B(\mu \to e\gamma) < 4.2 \times 10^{-13}$ [63]. The $l \to 3l'$ transitions have also been searched for at low energy τ factories such as Belle [64]. Currently no lepton flavour violating decays have been observed in any such decays. While these are very strong limits for many models, specific models regarding the Higgs boson can be better constrained by searches for direct LFV couplings.

In several LFV models, the addition of another Higgs boson-like doublet introduces the possibility of lepton flavour mixing, induced by a Yukawa coupling. A generic parameterisation [65] of lepton-flavour violating (LFV) couplings arising from Higgs couplings can be written as

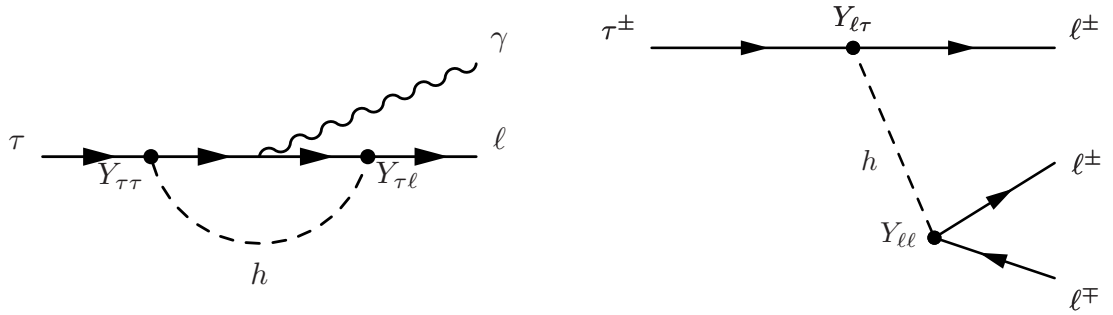
$$\mathcal{L}_{lep} \supset \sum_i (m_{e,i}\overline{e}_{L,i}e_{R,i} + \sum_j Y_{i,j}\overline{e}_{L,i}e_{R,j}h) + h.c. \tag{2.15}$$

where $Y_{i,j}$ are off-diagonal elements of a lepton mixing matrix[5] which couple a Higgs boson to two leptons of different flavours.

Searches for $H \to \ell\tau$ (at the LHC) , $\tau \to \ell\gamma$ and $\tau \to 3\ell$ (at Belle), where $\ell \neq \tau$, have sensitivity to these off-diagonal elements. $H \to \ell\tau$ decays have direct couplings but diagrams in figure 2.9 are how sensitivity to $\sqrt{|Y_{\ell,\ell'}|^2 + |Y_{\ell',\ell}|^2}$ enters for $\tau \to \ell\gamma$ and $\tau \to 3\ell$ decays. $H \to \ell\tau$ decays directly probe the coupling, rather than being a loop level coupling

[5]The diagonal component are the $H \to \ell\ell$ Yukawa elements

or by a factor of $Y_{\ell,\ell}$, respectively. Table 2.3 the upper limits of $\sqrt{|Y_{\ell,\ell'}|^2 + |Y_{\ell',\ell}|^2}$ (where $\ell \neq \ell'$) with a wide array of measurements.



(a) $\tau \to \ell\gamma$ (b) $\tau \to \ell^{\pm}\ell^{\mp}\ell^{\pm}$

Figure 2.9: Feynman diagrams of Higgs boson mediated LFV decays of τ leptons. The Yukawa coupling elements which are relevant to the matrix element calculation are noted on the vertices.

Decay process	Upper limits	Citation
Decays sensitive to $\sqrt{\|Y_{e,\mu}\|^2 + \|Y_{\mu,e}\|^2}$		
$H \to \mu e$	5.4×10^{-4}	[66]
$\mu \to e\gamma$	2.1×10^{-6}	[63]
$\mu \to eee$	3.4×10^{-5}	[67]
Decays sensitive to $\sqrt{\|Y_{e,\tau}\|^2 + \|Y_{\tau,e}\|^2}$		
$H \to \tau e$	2.3×10^{-3}	[68]
$\tau \to e\gamma$	1.4×10^{-2}	[69]
$\tau \to eee$	1.2×10^{-1}	[64]
Decays sensitive to $\sqrt{\|Y_{\mu,\tau}\|^2 + \|Y_{\tau,\mu}\|^2}$		
$H \to \mu\tau$	1.4×10^{-3}	[68]
$\tau \to \mu\gamma$	1.6×10^{-2}	[69]
$\tau \to \mu\mu\mu$	2.5×10^{-1}	[64]

Table 2.3: Current upper limits on LFV couplings to a Higgs boson from various decay modes [65, 70].

This table also shows the search with direct $H \to \mu e$ decays is much less effective than current $\mu \to e\gamma$ for measuring the $Y_{e\mu}$ element. The precision for $\mu \to e\gamma$ measurements far exceeds that of $H \to \mu e$ at the LHC by two orders of magnitude. Conversely, however,

searches for $H \to \ell\tau$ decays outperform the current $\tau \to \ell\gamma$ and $\tau \to 3\ell$ searches. Thus only searches for the $H \to \mu\tau$ and $H \to e\tau$ decays will be of interest for the LHC.

The measured sensitivity to $Y_{e,\tau}$ and $Y_{\mu,\tau}$ elements by ATLAS and CMS experiments [68, 71, 72] already exceed that of traditional $\tau \to \ell'\gamma$ and $\tau \to 3\ell'$ searches. Analysis of $\sqrt{s} = 8$ TeV data exhibited hints of a significant excess in search for $H \to \mu\tau$ [71, 73] decays. A 2.4σ excess over the SM expectation by CMS [73] and one signal region ATLAS observed a 2.2σ excess [71]. No such excess was found in the $H \to e\tau$ searches [66, 72]. This is expected due to the strong constraints from $mu \to e\gamma$ [63] which prevent the presence of two LFV Higgs boson couplings.

In an analysis of 35.9 fb^{-1} of $\sqrt{s} = 13$ TeV data [68], CMS found no excesses in any channel. The upper limits for the branching fraction of LFV $H \to \tau\mu$ and $H \to \tau e$ decays is 0.25% and 0.61% respectively, set by the CMS search at $\sqrt{s} = 13$ TeV. These limits also drive the upper limits for $\sqrt{|Y_{e,\tau}|^2 + |Y_{\tau,e}|^2}$ and $\sqrt{|Y_{\mu,\tau}|^2 + |Y_{\tau,\mu}|^2}$, placed at $< 2.26 \times 10^{-3}$ and $< 1.43 \times 10^{-3}$ respectively.

The goal of chapter 6 is to search for LFV $H \to e\tau$ and $H \to \mu\tau$ decays in $\sqrt{s} = 13$ TeV collision data with ATLAS. This search will aim to either confirm the excesses found in $\sqrt{s} = 8$ TeV data, or set even stronger limits on the LFV Higgs boson decays to τ leptons and the corresponding $Y_{i,j}$.

3 Experimental Setup

In order to study the Higgs couplings, a sufficiently large sample of Higgs boson decays is required. Since 2010, the high production rate of Higgs bosons at the Large Hadron collider (LHC [8]) has enabled the discovery of the Higgs boson in 2012 with the ATLAS and CMS experiments, as well as measurements of its numerous couplings. The LHC along with ATLAS [9] and CMS [74] detectors remain the only means of producing and detecting the Higgs boson.

The following sections will detail the experimental setup of the LHC as well as the ATLAS detector.

3.1 The Large Hadron Collider

The LHC is a particle collider situated on the border of Switzerland and France. The LHC can accelerate both proton beams and heavy ions, such as lead. The primary purpose of the LHC is to proton-proton (pp) collider. Its design allows for collisions of protons at centre of mass energies extending up to 14 TeV but is currently limited to 13 TeV collisions as of 2019.

Along the collider sits four experiments: ATLAS, CMS, LHCb [75] and ALICE [76]. The ATLAS and CMS detectors are general purpose detectors designed to study a wide range of physics phenomena including precise SM measurements as well as searches for new physics phenomena. Though the ATLAS and CMS detectors are similar in the physics goals, the detector designs have key differences. The LHCb detector is designed to study decays of B mesons, providing insight into flavour physics. The ALICE detector is principally used to study collisions of heavy ions with the purpose of studying quark-gluon plasmas.

3.1.1 Acceleration Process

The acceleration chain begins with the dissociation of gaseous hydrogen followed by the ionisation of hydrogen into protons. The protons are subsequently injected into the LINAC2 accelerator which increases the energy of protons to 50 MeV. Subsequently protons are injected to three sets of accelerators where the protons are further boosted. The protons move

from LINAC to the Proton Synchrotron Booster (PSB) to Proton Synchrotron (PS) and then to the Super Proton Synchrotron (SPS) where the beam energy goes from 50 MeV to 1.4 GeV to 26 GeV and then to 450 GeV.

After acceleration, the protons are injected into the main LHC ring as a beam of proton bunches. The two counter-rotating proton beams have their energies increased further via superconducting radio-frequency cavities to centre of mass energies as high as 14 TeV. As of 2018 the LHC has provided datasets corresponding to collisions at $\sqrt{s} = 7, 8, 13$ TeV. The beams are then focused to collision points located at the four detectors using quadrapole magnets. A schematic of the LHC is presented in figure 3.1.

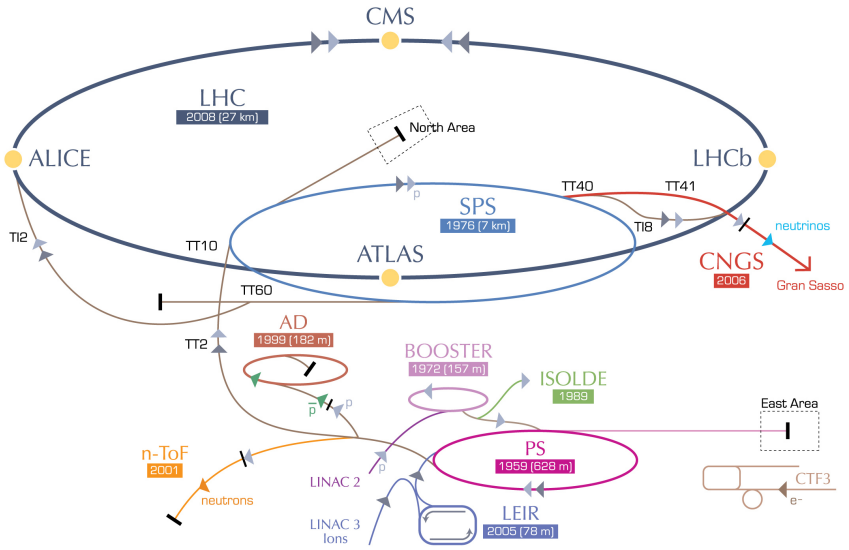


Figure 3.1: Schematic of the LHC apparatus and the associated experiments. Adapted from [77].

3.1.2 Proton Collisions and Beam Conditions

The number of events expected from collisions scales according to the luminosity of the beams and the cross-section of a physics processes. The former is parameterised by properties of the beams whereas the latter is dependent on the particular physics of interest. The instantaneous luminosity is a measure of the rate of collisions and is described by:

$$L = f_{\text{coll}} \frac{n_1 n_2}{4\pi\sigma_x\sigma_y} \tag{3.1}$$

where f_{coll} is the frequency of collisions, n_i are the number of protons in each bunch crossing and σ_i encode the beam profile. The total luminosity integrated over time measures the total of the number of collisions for a given dataset.

The LHC is designed to deliver an instantaneous luminosity of $1 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$, although currently the LHC is exceeding this. A planned upgrade is designed to deliver an even greater instantaneous luminosity after the completion of the current $\sqrt{s} = 13$ TeV collisions. A summary of collision parameters is reported in table 3.1, The total integrated luminosity as a function of time given in figure 3.2.

The datasets used in this thesis consist of data collected in proton-proton collisions in 2015 and 2016 at $\sqrt{s} = 13$ TeV. The total integrated luminosity of these datasets is 36.1 fb^{-1} unless otherwise stated.

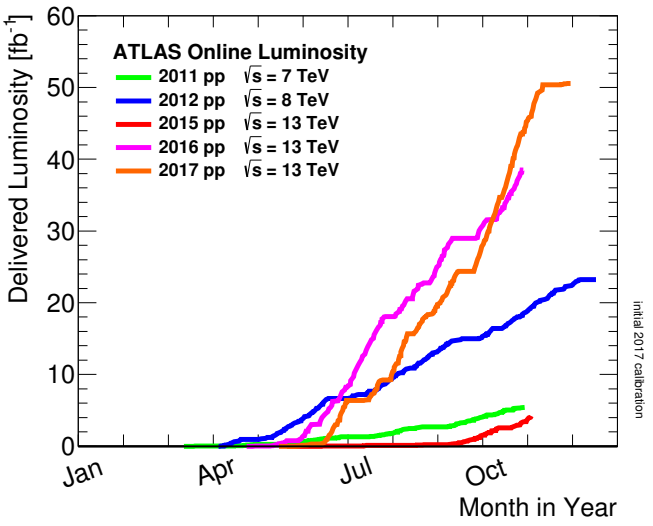


Figure 3.2: Luminosity delivered by the LHC across the years the LHC was active as measured by the ATLAS experiment [78].

Beam Conditions	2015	2016
$\sqrt{s}$ [TeV]	13	13
Bunch spacing [ns]	25	25
Typical Bunch Population [10^{11} protons/bunch]	1.2	1.2
Peak luminosity [$10^{33} cm^{-2} s^{-1}$]	6.3	11
Peak number of inelastic interactions per crossing	40	39
Average number of interactions per crossing	21	27
Total Integrated Luminosity Delivered [fb^{-1}]	5.0	31

Table 3.1: Selected parameters describing the beam conditions delivered by the LHC during data-taking for datasets used here. [79, 80].

3.1.3 Pileup

Nominally, events are recorded in the instance where inelastic collisions occur between partons, either gluons or quarks. These collisions are referred to as "hard scatterings" and create the physics phenomena of interest. In a typical collision at the LHC however, there may exist several interactions whenever bunches of protons cross at the interaction point.

This phenomena is known as "in-time pileup" (commonly just pileup). The nature of these multiple secondary interactions is of "softer" elastic collisions but create issues in discerning the origin of particles within a recorded event due to the frequency at which this occurs. Large rates of pileup are detrimental to the reconstruction of both events and particles.

The "out-of-time pileup" results from limitations of the readout electronics compared to the timing bunch crossing spacing. This occurs when events are still being recorded while a second bunch crossing has occurred.

The key metric for the amount of pileup is the number of interactions per bunch crossing, called μ. The distribution of the average μ ($< \mu >$) of recorded events for the datasets used in this thesis are presented in figure 3.3. The difference between 2015 and 2016 is due to an increase in the instantaneous luminosity.

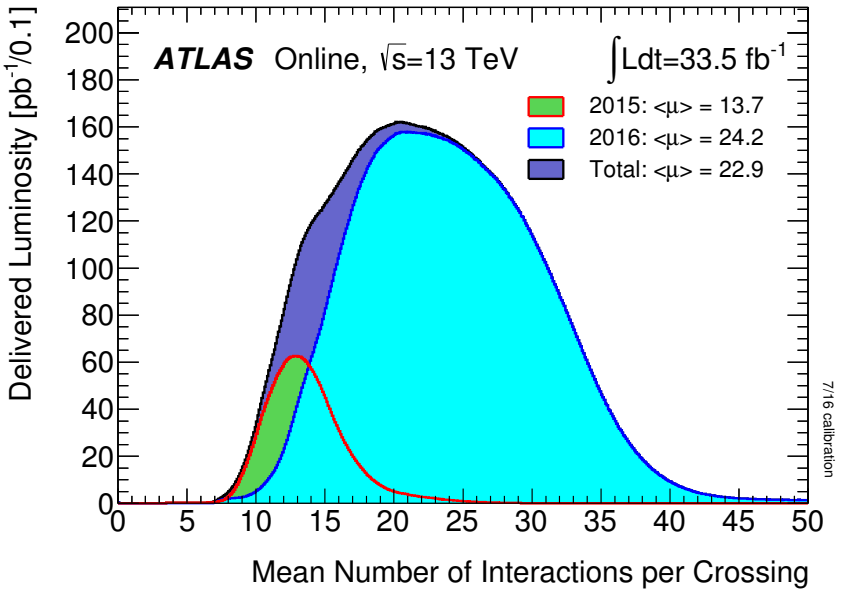


Figure 3.3: Luminosity as a function of the average number of interactions per bunch crossing [78].

3.2 The ATLAS Detector

The ATLAS detector, was designed to perform measurements of SM and searches for diverse range of new physics in parallel with the CMS detector. The following sections describe the ATLAS detector and how the particles are reconstructed for each event.

3.2.1 Coordinate System of ATLAS

The collisions which produce hard scattering events at the LHC occur between partons. These partons do not have a known energy as they carry an unknown fraction of the total centre of mass energy of the proton bunches. While this implies the centre of mass energy of any given collision is unknown, what is known is there is little momentum transverse to the beam direction. Hence a co-ordinate system which is invariant under boosts along the beam axis as well as symmetric about the beam axis.

A right-handed Cartesian coordinate system is defined with the interaction point as the origin. The z direction is in the direction of the beam, the x axis points toward the centre of the LHC ring and the y axis points in the upward direction to the surface. The kinematics of a given particle candidate are described by the following quantities:

- p_T- the momentum transverse to the z axis
- ϕ - the azimuthal angle in the transverse x-y plane
- y - the "rapidity" which depends on the longitudinal (along the beam) kinematics

The rapidity is defined as:

$$y = \frac{1}{2} \log \frac{E + p_L}{E - p_L} \tag{3.2}$$

with p_L being the longitudinal component of the momenta. In the limit of $E >> m$, which holds for most particle candidates at the LHC, the pseudorapidity (η) is used instead:

$$\eta = -\log \left[\tan \left(\frac{\theta}{2} \right) \right] \tag{3.3}$$

where θ is the corresponding polar angle with respect to the z direction. This quantity is simpler to measure as it depends only on the θ angle. Thus the momenta is given with p_T, ϕ and η as its primary constituents.

The typical distance measure is:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{3.4}$$

For the purpose of tracking, the measure of impact parameters is invariant under Lorentz boosts along the z direction and are thus insensitive to the longitudinal momenta of the initial

state partons. In this case the system includes z_0 and d_0, the point of closest approach in the z-axis and the x-y plane are added. The total coordinate system of tracking is illustrated in 3.4.

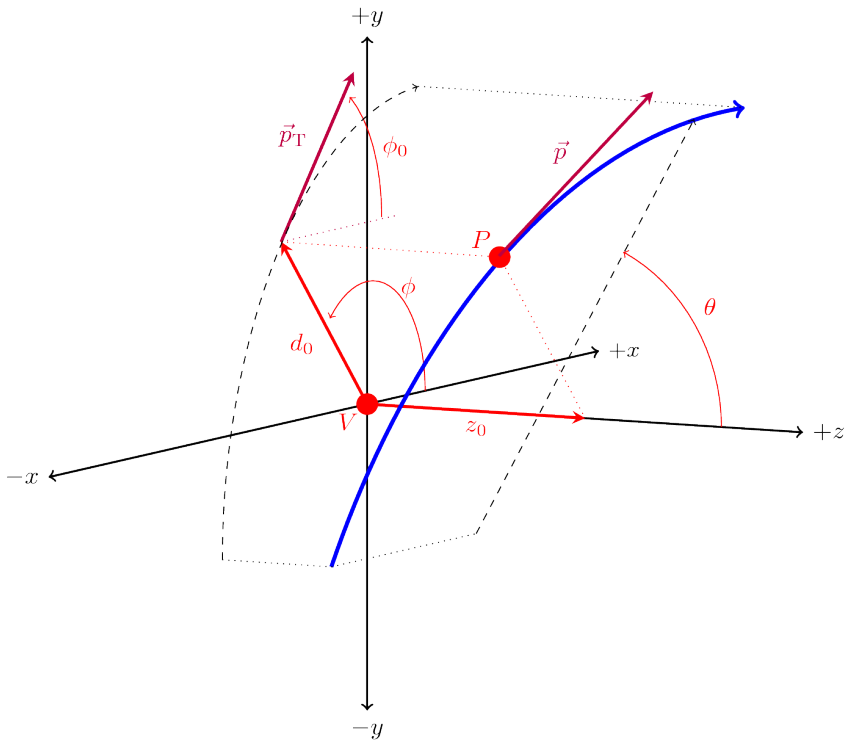


Figure 3.4: The definition of angles for the ATLAS coordinate system.

3.2.2 Detector description

The detector is divided into three main sections, the inner detector, the calorimeter system and the muon spectrometers. Each detector component serves a different purpose in the reconstruction of particle candidates. A schematic of the ATLAS detector is offered in figure 3.5 and the subsequent design resolution and η coverage is summarised in table 3.2. Each detector component is separated into two configurations. The central (barrel) region have layers running parallel to the beampipe and the forward (endcap) region have layers running perpendicular to the beampipe.

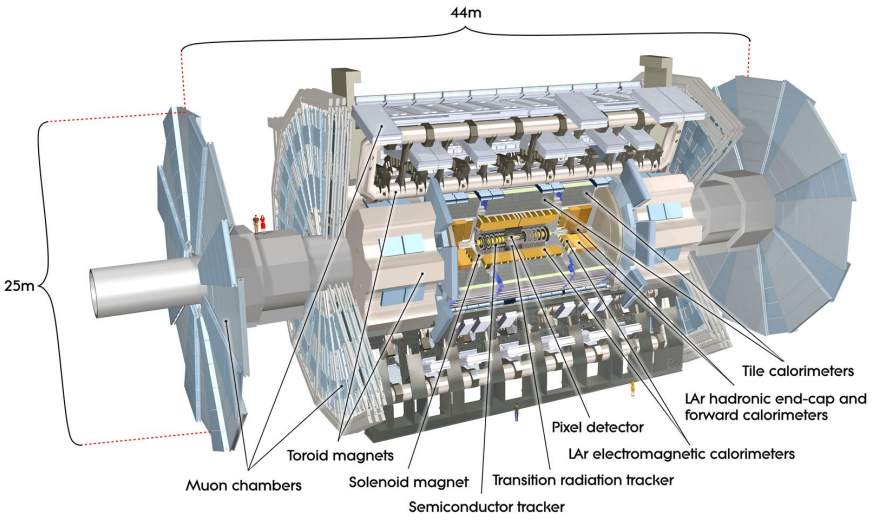


Figure 3.5: A cross section of the ATLAS detector [9].

The ATLAS detector aims to be a general-purpose detector which can handle the intense radiation environments of the LHC. Some of the keys to its effectiveness are:

- Electronic components which have low latency and are radiation resistant.
- High granularity detector components which can handle large particle fluxes and distinguish overlapping events.
- Large geometric acceptance in both angular directions.
- Fine resolution for charged particle momenta with excellent reconstruction efficiency via the inner detector.
- An electromagnetic calorimeter system for measuring electrons and photons with precision along with a hadronic calorimeter for jet measurements. The combined calorimeters allow for E_T^{miss} measurements.
- An accurate muon spectrometer allowing for precise momenta resolution and precise charge reconstruction.
- Efficient trigger systems which are both fast and reliably select interesting physics events.

3.2.2.1 Inner Detector

The inner detector is dedicated to the tracking of charged particles. Located closest to the beampipe, it is encased by a solenoidal magnet producing a 2T field used to bend charged

Detector component	Target resolution	$\lvert\eta\rvert$ coverage	
		Measurement	Trigger
Tracking	$\frac{\sigma_{pT}}{p_T} = 0.05\%\ p_T \oplus 1\%$	< 2.5	-
Electromagnetic calorimetry	$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%$	< 3.2	< 2.5
Hadronic calorimetry			
Barrel and endcap	$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$	< 3.2	< 3.2
Forward	$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\frac{\sigma_{pT}}{p_T} = 10\%$ at $p_T = 1$ TeV	< 2.7	< 2.7

Table 3.2: Performance goals of the ATLAS detector [9]. Note that for high p_T muons, the muon spectrometer performance is independent of the inner detector system. E and p_T are in GeV unless otherwise specified.

particles. The trajectory is measured when the charged particle interacts with elements of the inner detector, leaving energy deposits re-clustered as "hits". The coverage of the inner detector is covers the full azimuthal direction and extends to $|\eta| < 2.5$ in the polar direction. The inner detector is subdivided into a number of subcomponents. A schematic of the layering of these components is presented in figure 3.6.

Pixel

The pixel detector is located closest to the beampipe and is designed to provide the finest resolution measurements. Since 2015, an insertable b-layer (IBL), originally not included in the initial commissioning of the ATLAS detector, has been integrated to deal with the vast increase in the number of charged tracks expected in $\sqrt{s} = 13$ TeV collisions. The intrinsic resolution of this layer is 10 μm in $(r - \phi)$ and 115 μm in z.

Semiconductor Tracker (SCT)

The following layer from the pixel is the SCT. The SCT is constructed in four layers of strips cylindrically arranged about the beampipe with eight additional disc layers for the end-caps. The intrinsic accuracy of this component is 17 μm in $(r - \phi)$ and 580 μm in z.

Transition radiation detector (TRT)

The TRT is designed to capture tracking information using transition radiation rather than direct hits. As charged particles enter the 4mm straw tubes of a Xenon gas mixture, the gas ionises and electrons drift toward the charged wire in the centre of the tube. The resolution

in the $(r - \phi)$ direction is 130 μm within the barrel region of $|\eta| < 2$. No measurement in the z direction is available.

The TRT's inclusion is designed to improve particle identification, particularly for electrons. The different masses of particles interacting with the TRT produce specific transition radiation, due to passing between materials of different refractive indexes, at different momenta.

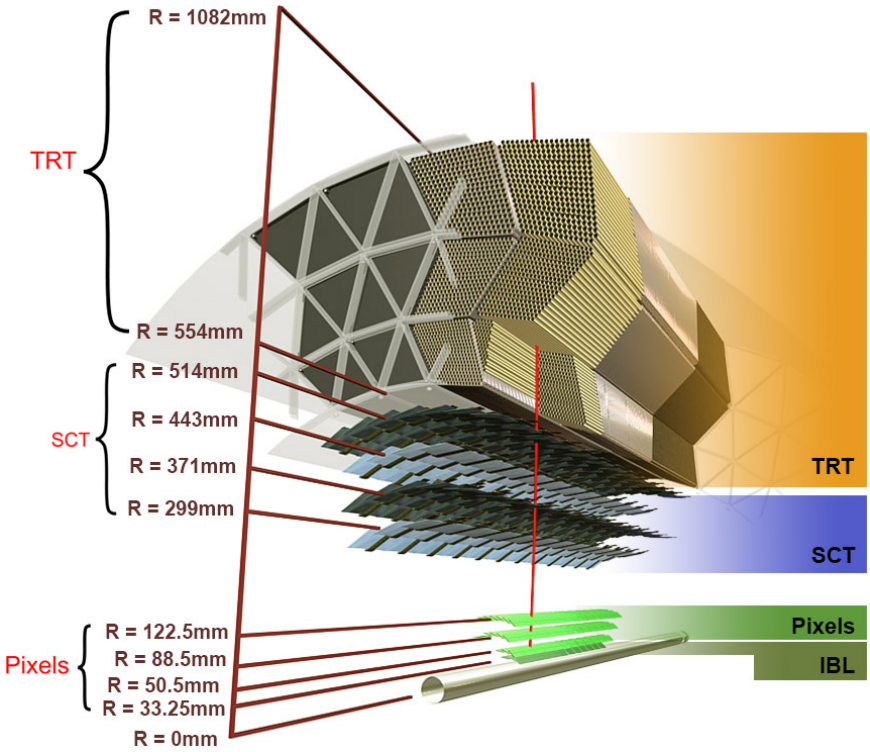


Figure 3.6: A cross section of the ATLAS inner detector [9].

3.2.2.2 Calorimeter System

The calorimeter system contains the bulk of the material in the ATLAS detector and is used to measure energy deposits of electrons, photons, taus and jets. As particles pass through the calorimeter systems, interactions with the calorimeter induce decays of particles leaving deposits of energy, ideally until the entire energy of the particle is absorbed. Each calorimeter subcomponent consists of a barrel and forward section for each of the electromagnetic and hadronic calorimeter detector systems. A cross-sectional view of the calorimeter system is illustrated in figure 3.7.

Electromagnetic Calorimeter (ECAL) The ECAL's primary purpose is to absorb and measure the energies of electrons and photons. The ECAL is constructed of liquid argon (LAr) with a thickness exceeding 22 and 24 radiation lengths for the barrel and end-caps

respectively to fully capture the cascade decay of the electron/photon. The shape of the ECAL is designed in an accordion shape of three-four layers, in order to provide full coverage in the azimuthal direction. The ECAL's barrel and end-cap components extend coverage to $|\eta| < 1.475$ and $1.375 < |\eta| < 3.2$ respectively. The ECAL is the finer granularity of the two components with 0.025×0.025 in $\Delta\eta \times \Delta\phi$ in the barrel region to 0.1×0.1 in the end-cap.

Hadronic Calorimeter (HCAL) The thicker HCAL is designed as the final layer to stop all particles with the exception of muons and neutrinos. As the name suggests the HCAL focuses on hadronic activity comprising mostly of jets.

The barrel component is referred to as the tile calorimeter and has a thickness of 9.7 interaction lengths and has a range of $|\eta| < 1.7$ with a granularity ranging from 0.1×0.1 to 0.2×0.1. The tile component is made of scintillating tiles and a steel absorber. The scintillating material emits photons upon interaction with an ionising particle which is amplified with photo-multiplier tubes. The final steel layer prevents punch-through from remaining energetic hadrons.

The LAr endcaps capture particles in the range to $3.1 < |\eta| < 4.9$ and an additional forward LAr detector extends the range to $3.1 < |\eta| < 4.9$ capturing very forward jets. The forward LAr detector is comprised of a copper layer which is used for electromagnetic interactions followed by two tungsten layers to measure hadronic interactions. The overall endcap/forward region has an effective thickness of 10 interaction lengths.

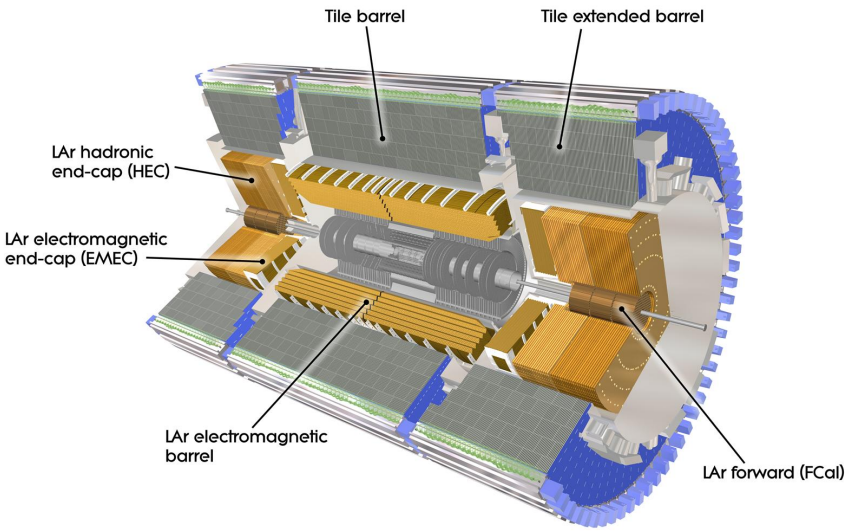


Figure 3.7: A cross section of the calorimeter system of the ATLAS detector [9].

3.2.2.3 Muon Spectrometer

Muons, unlike most other particles detected, are typically minimally ionising which results in them passing through the calorimeter without leaving energy depositions. Thus a dedicated system was created, designed solely to detect and measure muons. The muon spectrometer, the furthest component of the ATLAS detector from the interaction point, consists of the tracking/trigger systems and the magnets.

A set of air-core toroidal magnets are placed outside the calorimeter system which bends the muons passing through the magnetic field. In the core barrel region $|\eta| < 1.4$ the toroid can produce up to 5.5 Tm of bending power[1] and up to 7.5 Tm from the end-caps which extends out to $|\eta| < 2.7$.

Beyond the toroid magnets lies the final layer of detectors. The Monitored Drift Tubes (MDTs) and Cathode Strip Chambers (CSCs) are responsible for the tracking of the muons. The MDTs are filled with an Ar/CO_2 and cover a range up to $|\eta| < 2.7$ whereas the CSCs are multiwire chambers covering specifically the $2.0 < |\eta| < 2.7$ region where high rates requires a more specialised equipment.

The triggering system relies upon the Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs) that cover ranges of $|\eta| < 1.05$ and $1.05 < |\eta| < 2.7$, respectively. In addition to providing clear p_T thresholds, the muon chambers provide good measurement of the azimuthal coordinate and aids in the identification of bunch crossings.

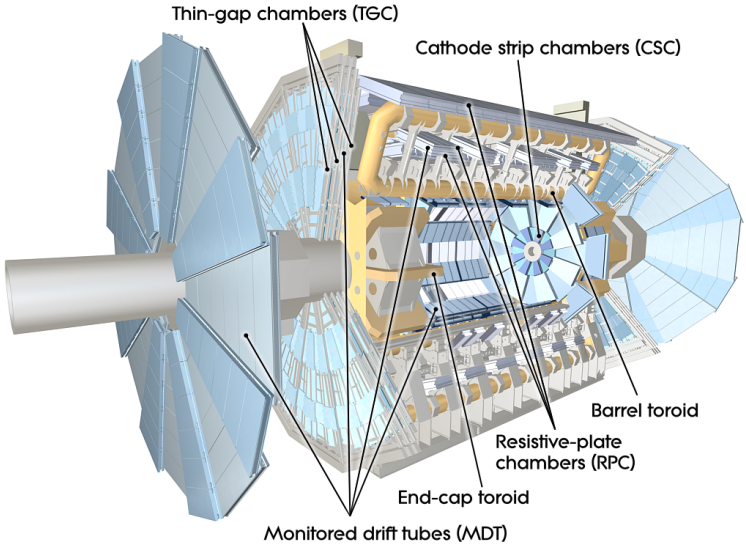


Figure 3.8: A cross section of the muon spectrometer of the ATLAS detector [9].

[1] the magnetic field strength integrated over the path

3.2.3 Trigger and Data Acquisition

With a bunch spacing between proton bunches of 25ns, the rate of collisions is 40 MHz. Due to the limited data-recording rate of approximately 200 Hz, only a subset of all collisions can be recorded. To effectively select which events are stored, a multi-staged trigger and data acquisition (TDAQ) system is implemented.

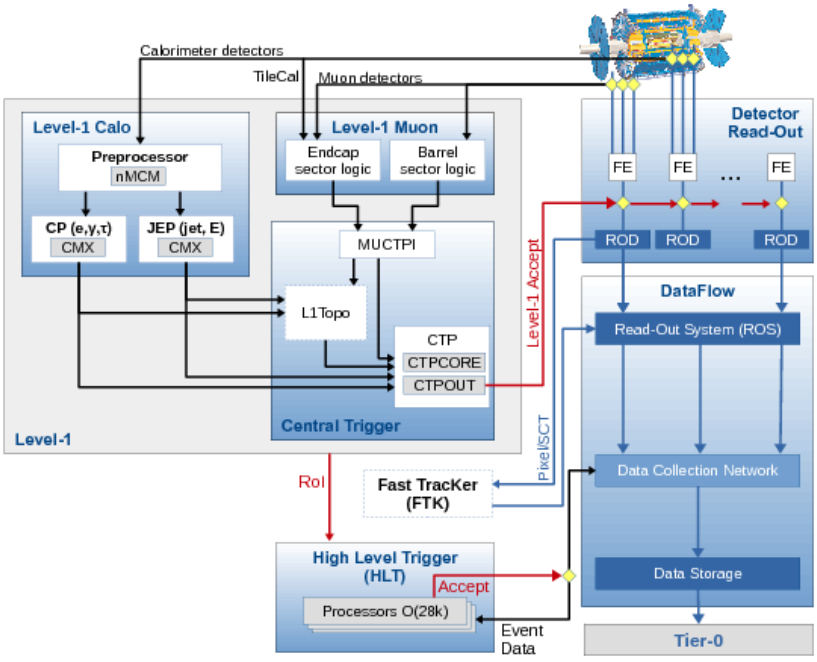


Figure 3.9: A schematic of the ATLAS trigger system as of 2015 [81].

The first stage of the TDAQ system is the Level-1 (L1) hardware trigger, which aims to reduce the rate to 70kHz. The L1 triggers are based on the major detector components of the calorimeter (L1Calo) and muon spectrometer (L1Muon) as well as a topological based trigger (L1Topo). The L1Calo and L1Muon triggers make broad selections on very simple reconstructed variables (such as p_T) in order to reduce this rate. The L1Topo calculates and selects based on a few simple topological variables such as the ΔR or masses between particles.

After the rate is reduced to a manageable amount by the L1 system, information is passed to the "High Level Trigger" which is the final stage of the trigger system. The HLT is a software based trigger which utilises information not accessible by the L1 hardware-based trigger. More refined information from the calorimeter, inner detector and muon spectrometers are used with a finer granularity. High level variables such as identification scores, isolation and missing energy can be also be reconstructed in this step. The HLT triggers

finally reduce the rate to that which can be recorded to disk.

HLT triggers can be defined for use in broad or very specific physics cases. For the studies shown in later chapters, triggers will involve the selection of a single electron or muon passing p_T thresholds and identification requirements. While it is possible to trigger on many other signatures for the studies presented, the single electron and muon triggers have the advantage of a relatively low p_T threshold as well as a high reconstruction efficiency.

3.2.4 Particle Candidate Reconstruction

Particles are reconstructed by combining information from different components of the ATLAS detector. Brief descriptions of leptons, photons, jets and missing energy are provided here. The reconstruction of τ candidates is detailed in chapter 4.

3.2.4.1 Electrons and Photons

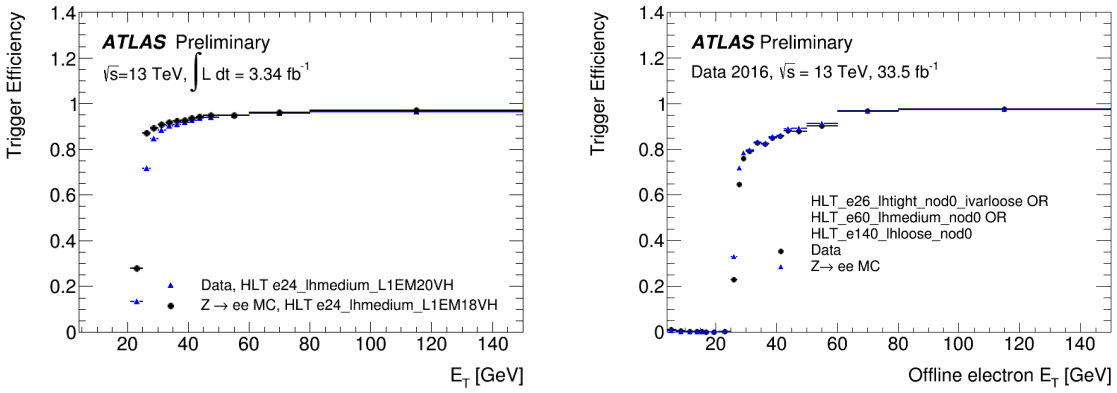
Electrons are reconstructed combining information from tracking components with calorimeter deposits in the electromagnetic calorimeter. Tracks with p_T exceeding 0.5 GeV within the acceptance of the inner detector are extrapolated toward the centre of the ECAL and matched to clusters using a sliding window to compare seeds. No candidates are considered within the crack in the ATLAS calorimeter of $1.37 < \eta < 1.52$. Candidates are accepted within a fiducial region of $E_T > 15$ GeV and $|\eta| < 2.47$.

Once electron candidates are reconstructed, identification requirements are imposed to reject candidates originating from background sources such as jets and photon conversions. The identification algorithm used for electrons is a multivariate classifier score based on likelihoods of input variables. The input variables include information collected by the tracking detectors as well as EM calorimeter shower shapes. Three working points are defined with varying signal efficiency targets: loose, medium and tight. The loose working point is used to define electron candidates, with an efficiency of at least 90%, unless otherwise indicated [82].

In addition to the identification, further stringent isolation requirements can be applied to reject candidates with either an excess of calorimeter deposits or additional tracks in a cone of $\Delta R = 0.2$ about the electron candidate. Several working points are defined but only the *gradient* isolation will be further considered. This working point has an efficiency of 90% and 99% with $E_T = 20$ and $E_T = 60$ GeV respectively [82].

Due to the high efficiency of electron signature, the electron trigger is one of the primary triggers used henceforth. The single electron trigger utilises L1 calorimeter based seeds and at HLT requires a lower limit on the p_T with triggers used for the 2016 data-taking period re-

quiring passing certain identification requirements. During 2016, isolation requirements are necessary due to increases in rates due to an increase in presence of pileup. For simplicity, a gradient isolation is required regardless of dataset.[2] The efficiency curves are presented in figure 3.10.



(a) 2015 [83] (b) 2016 [84]

Figure 3.10: Efficiency curves for the lowest usable single electron HLT triggers in data for two data-taking periods.

Corrections for differences in efficiencies of the reconstruction, identification, isolation and trigger between MC and data are derived using $Z \to ee$ and $J/\psi \to ee$. The method used to select events, the "tag and probe method", is very similar to those outlined in chapter 4 except using a different signature to focus on the electron candidates [82].

Photons are reconstructed similarly to electrons but leave no tracks in the inner detector. Photons will not be considered further as no photons are present in any of the signatures being discussed.

3.2.4.2 Muons

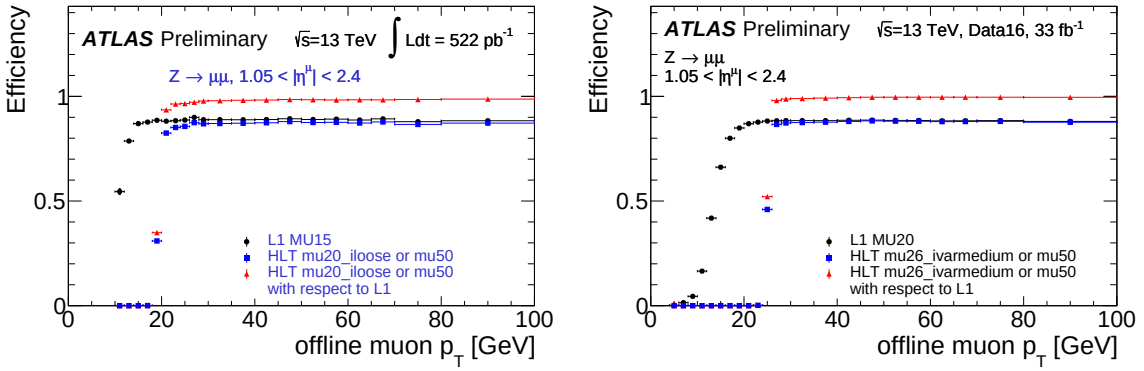
Muons are reconstructed from hits in the inner detector and the muon spectrometer using a global fits χ^2 as a selection criteria. Several reconstruction algorithms can be defined as muons depending on which components of the detector are used. For the measurements to be described only the "combined" algorithm, where tracking is done independently in inner detector and muon spectrometer and then combined, is used for final candidate selection [85].

[2]Some background estimates require an inversion of this requirement. This will be explicitly stated when required.

Identification algorithms focus on track quality and a charge over momentum (q/p) ratio. These requirements aim to surpress long lived pions and kaons. The loosest requirement has a selection efficiency of 97.5% in the barrel region, and utilises muons of any reconstruction [85]. The medium identification requirement calls for a minimum of three hits in the muon spectrometer with two of those originating from the MDT. A further requirement on the significance of the q/p ratio is required. The main fiducial acceptance of muons is within $|\eta|< 2.5$ with a minimum threshold of $p_T > 10$ GeV. The overall reconstruction efficiency exceeds 95% at this working point [85]. Muons passing a medium requirement will be used for final candidate selection whereas the loose requirement will be used to remove other particle candidates which geometrically overlap.

Like electrons, isolation requirements are also utilised with the same efficiency working point being employed. A wider cone of $\Delta R = 0.3$ is used for track isolation compared with isolation requirement for electrons. Tag and probe techniques are used on high purity $Z \to \mu\mu$ and $J/\psi \to ee$ events to derive corrections for the identification and isolation requirements.

Similarly to the electron signature, the muon signature has high efficiency so will be used as one of the triggers for measurements detailed in the next chapters. The single muon triggers start with seeds from muon spectrometer with candidates required to have $p_T> 20$ GeV. The efficiency curves are presented in figure 3.11.



(a) 2015 [**mu˙trig**] (b) 2016 [86]

Figure 3.11: Efficiency curves for the lowest usable HLT triggers in data for two data-taking periods.

3.2.4.3 Jets

If a process produces a quark or gluon during hard scattering, colour confinement in QCD forces hadronisation of these particles. This process continues in a cascade until stable

colourless hadrons are formed. Due to the energy in the collisions, the hadronisation forms a complex shower of particles called a *jet*. Jets are the most common signature in the ATLAS detector, characterised primarily with calorimeter deposits.

Jets reconstruction is initiated by the identification of localised clusters of calorimeter deposits defined by the topoclustering algorithm [87]. These localised clusters are known as TopoClusters. The TopoClusters are now combined in the anti-k$_T$ algorithm with a distance parameter of ΔR= 0.4 [88]. Jets energy is calibrated at the electromagnetic energy scale initially and then recalibrated ultimately with a data-driven approach to the Jet Energy Scale (JES) [89]. For jets with $p_T < 50$ GeV, jets are required to pass a vertex tagging algorithm score of 0.4 and 0.6 for the forward [90] and barrel [91] regions respectively. Henceforth only jets with $p_T > 20$ GeV in a fiducial region of $|\eta| < 4.5$ will be considered.

As the rate of b quark production is extremely high in collisions, specialised identification is utilised for discrimination of b quark-initiated jets, known commonly b-jets. A key characteristic of b-jets in comparison to those initiated from lighter particles is the fact that b-jets will have longer decay lengths. As the primary decay to b-hadrons is rather stable, the b-jets is expected to produce a secondary vertex which can be easily detected. A multivariate classifier combines information about impact parameter and secondary vertex in order to reject jets originating from lighter quark decays. A working point of 85% efficiency is defined to select b-jets from top pair production [92].

3.2.4.4 Missing Transverse Energy

Missing energy is often present due to the presence of neutrinos produced in hard scatterings. As the centre of mass energy is not known, only the transverse component of the missing energy vector is relevant. Though not actually a particle, the reconstruction of missing transverse energy (E_T^{miss}) is vital in processes involving neutrinos.

The E_T^{miss} can be expressed as a sum of the transverse components of other particle candidates:

$$E_T^{miss} = -\sum_i E_T^i = -(E_T^{jets} + E_T^{electrons} + E_T^{\mu} + E_T^{\tau} + E_T^{\gamma} + E_T^{soft}) \tag{3.5}$$

where E_T^i represents the vector component in the transverse plane of reconstructed particle i which is associated with a vertex. The E_T^{soft} term refers to tracks or calorimeter clusters which do not pass the reconstruction requirements detailed prior. The E_T^{miss} is a variable which is rather sensitive to the reconstruction of the event as a whole and has a generally poor reconstruction resolution in comparison to other particle candidates [93].

3.2.5 Monte-Carlo Simulations

To model the backgrounds and signals expected at the LHC, simulations are primarily used. These simulations use a Monte-Carlo method of random sampling of a probability distribution function and are thus known as Monte-Carlo simulations (MC). The simulation of events involves a number of steps: the calculation of matrix elements, decay and showering of unstable particles and finally the detector simulation.

The modelling of the initial collision is particularly difficult. The partons which collide in the proton beams carry only a fraction of the energy of the proton. The distribution of the expected fraction of energy carried is described by a parton distribution function, which is experimentally determined. Many different parton distribution functions are available to be interfaced with the calculation of matrix elements and involve experimental data from LHC collisions.

After the generation of the scattering, the showering of the decay products, including initial state radiation, is performed. At this stage the presence of initial and final state radiation becomes important. The showering is performed by decaying partons at a certain level of accuracy in QCD in a continual process until calculations produce particles which are stable with energies at the hadronic cut-off scale of approximately 200 MeV. QED radiation such as bremsstrahlung is also accounted for in a similar manner.

The interaction of the final stable particles and showers with the ATLAS detector is done via a specialised geometry in GEANT4 [94]. The generators specific to particular studies will be detailed accordingly in the appropriate sections. As MC simulation is generated prior to collisions, pileup needs to be simulated. Pileup vertices are generated with Pythia and added atop the main hard scattering vertex. The pileup effects are then corrected after data is collected by a reweighting process based on correcting pileup dependent distributions such as the average $< \mu >$ profile and N_{vertex}.

4

Tau Reconstruction and Calibration

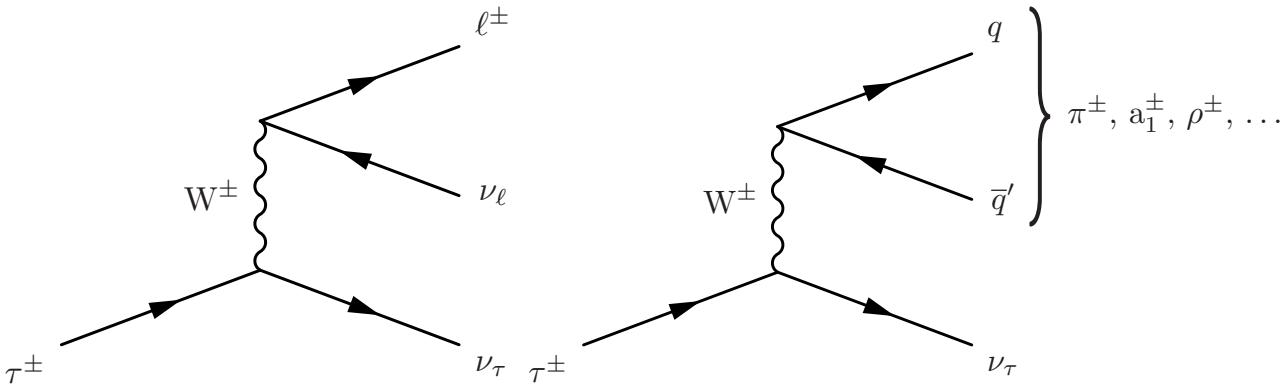
The central focus of this thesis is to measure SM or BSM Higgs boson couplings to τ leptons. In this chapter a discussion of τ lepton reconstruction is presented.

Of the three charged leptons, the τ lepton is the heaviest ($m_\tau = 1.78$ GeV) resulting in a very short lifetime ($c\tau \approx 85\mu$m). This presents a challenge in detecting the particle directly with the ATLAS detector. Instead, decay products of the τ lepton are reconstructed.

Being so heavy, the τ lepton has a number of different decay modes including to leptonic final states and hadronic final states. Hadronic final states, mostly charged and neutral pions with one neutrino, constitute a majority of the τ lepton decays ($\sim 65\%$) and are termed τ_{had}. These hadronic states often decay through resonant intermediate meson states such as $\rho^{\pm}(770)$ which decays to $\pi^{\pm}\pi^0$ and $a_1(1260)$ which decays to $\pi^{\pm}\rho^0$ to a final state of three pions. Leptonic final states ($\sim 35\%$) consist of one charged lepton - either electron or muon - and two neutrinos. Feynman diagrams depicting the decay modes of the τ lepton are presented in 4.1. A breakdown of the branching fractions is provided in figure 4.2.

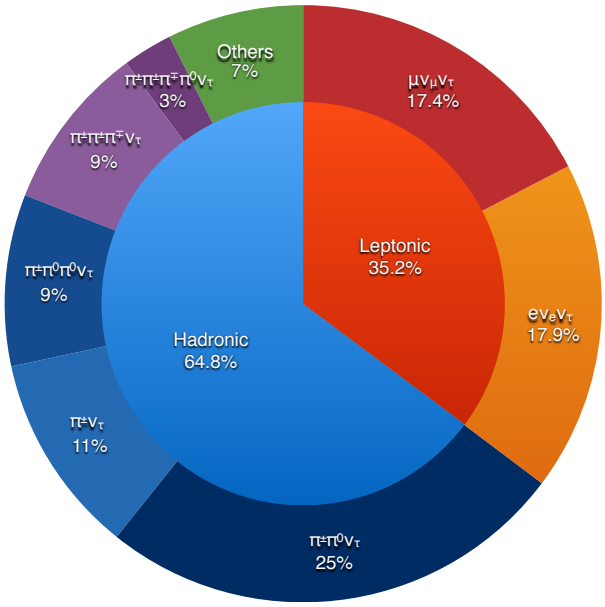
Leptonic decay modes are indistinguishable from electrons and muons produced in other processes in proton collisions. Conversely, the hadronic final state forms a distinctive collimated shower of particles which is easier to distinguish from other particle signatures upon interaction with the ATLAS detector. Hence a dedicated reconstruction, identification and calibration are developed specifically for τ_{had} candidates.

This chapter shall summarise how hadronically decaying τ leptons are reconstructed, identified and calibrated by the ATLAS experiment. Particular emphasis will be placed on the calibration of the energy of the hadronic decays of τ leptons using $Z \to \tau\tau$ decays.



(a) Leptonic Decay (b) Hadronic Decay

Figure 4.1: Feynman diagrams illustrating the two types of τ lepton decay.



Decay Mode	BR (%)
Leptonic	35.3
$e\nu_e\nu_\tau$	17.9
$\mu\nu_\mu\nu_\tau$	17.4
Hadronic	64.7
One prong	45.6
$\pi^{\pm}\nu_\tau$	10.8
$\pi^{\pm}pi^0\nu_\tau$ (via $\rho^{\pm}(770)$)	25.5
$\pi^{\pm}\pi^0\pi^0\nu_\tau$ (via $a_1^{\pm}(1260)$)	9.3
Three prong	12.0
$\pi^{\pm}\pi^{\pm}\pi^{\pm}\nu_\tau$ (via $a_1^{\pm}(1260)$)	9.3
$\pi^{\pm}\pi^{\pm}\pi^{\pm}\pi^0\nu_\tau$	2.7
Others	7.1

Figure 4.2: Branching fractions for decay modes of τ leptons [95]. Note the fractions in the table for one prong and three prong exclude Kaon channels.

4.1 Reconstruction

The signature of a τ_{had} has an experimental signature similar to that of a narrow jet. Thus τ_{had} reconstruction originates from selecting jet candidates reconstructed using the anti-k_T algorithm, with $\Delta R = 0.4$ as the distance parameter.[1] These jets seeds are calibrated using local hadronic calibration (LC) [96], which reconstructs energies at the particle level. This method compensates for various detector effects such as non-compensation of the detector, dead material and the energy of deposits outside the $\Delta R< 0.4$ cone. Only candidates within a fiducial region of $p_T > 10$ GeV and $|\eta| < 2.5$ are selected. Candidates reconstructed in the "crack region" of $|\eta| \in [1.37, 1.52]$ are excluded due to the degraded performance corresponding to the region where barrel and end-cap EM calorimeters meet.

Some quality track requirements are also applied to ensure a well constructed track:

- Tracks associated with the τ_{had} must have a reconstructed $p_T > 1$ GeV.
- Impact parameter requirements of $|d_0| < 1$ mm and $|\Delta z_0 \sin\theta| < 1.5$ mm.
- Track must also have a minimum of two hits in the Pixel detector.
- Seven total hits from the Pixel and SCT systems.

The vertex of the τ lepton decay is taken as the vertex with the largest fraction of track momenta associated amongst all reconstructed vertices. The central axis of the τ_{had} is taken as the barycenter of associated TopoClusters [87] and two cones are defined in order to later discriminate against QCD jets. The core cone is defined for clusters and tracks within $\Delta R < 0.2$ and the isolation cone is defined from $\Delta R \in [0.2, 0.4]$.

Finally τ_{had} candidates are separated based on whether they have one or three charged tracks (also known as prongs): henceforth these will be known as 1p and 3p respectively. The separation of candidates into 1p and 3p is largely so identification algorithms can be separately optimised for the unique composition of signatures which mimic the τ_{had}.

[1] See section 3.2.4.3 for a description of jet reconstruction

4.2 Identification

The reconstruction algorithm, summarised in the previous section, reconstructs τ_{had} from jet candidates. This would imply many jets, originating from the hadronisation of a gluon or quarks rather than a τ lepton decay, can be misreconstructed as a τ_{had}. These jets are termed "QCD jet" or simply "jet".

The cross-section of QCD jets which are initiated by the hadronisation of quarks and gluons dwarf that of the τ lepton production. Without a dedicated τ_{had} identification, reconstructed τ_{had} would consist almost entirely of QCD jets. In order to properly identify τ_{had}, a multivariate algorithm is developed to distinguish τ_{had} from QCD jets based on the complex decay topology of a τ_{had}.

Typically there are four key differences between QCD jets and τ_{had}:

1. QCD jets typically produce a wider shower of particles than a τ_{had} for a given p_T.
2. The charged pions in a τ_{had} carry a higher fraction of the visible energy.
3. The number of neutral pions is lower for τ_{had} candidates.
4. A τ_{had} typically feature a displaced vertex or a significant impact parameter, whereas this is not necessarily true for QCD jets.

These features motivate the selection of inputs variables which are combined in a Boosted Decision Tree (BDT), a supervised learning multivariate classifier.[2] The classifier is trained to distinguish signal, consisting of simulated $Z \to \tau\tau$ events, and background, which comprises simulated dijet events. The inputs are rescaled such that the mean of the variables for true τ_{had} candidates is constant with respect to the average $< \mu >$. The inputs are combined in two separate BDTs, one trained for events with 1p τ_{had} and the other with 3p τ_{had}, in order to separate true τ_{had} candidates from QCD jets. The output of the classifier is a score between 0 and 1. Sores closer to 0 indicate a QCD-like jet and scores closer to 1 indicate a τ_{had}. This is illustrated for the training on MC generated for collisions at $\sqrt{s} = 13$ TeV in figure 4.3.

[2] See [97] for more details. Further information on BDTs will be discussed in 6.4.1

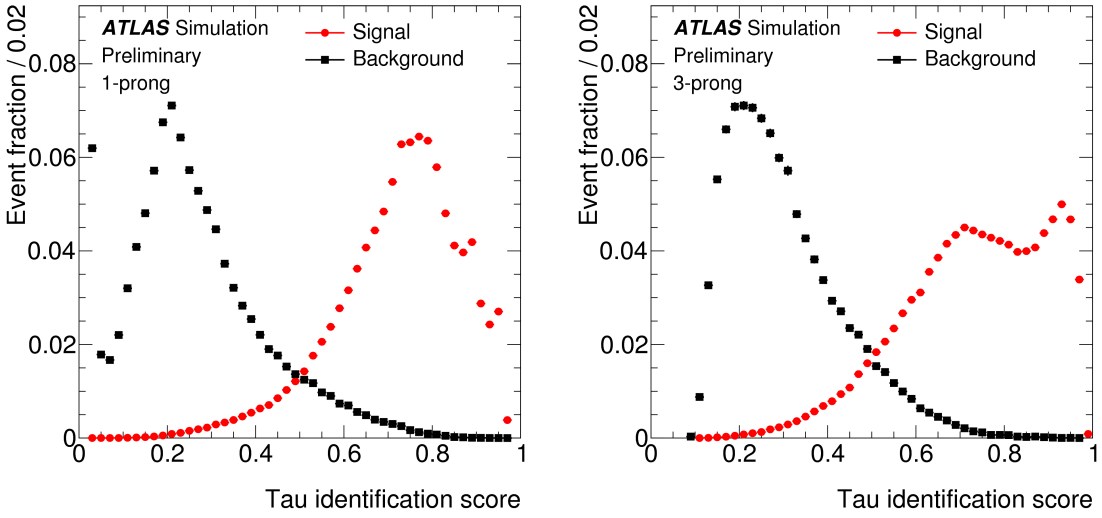


Figure 4.3: The response of the identification BDT for one (left) and three (right) pronged τ_{had} [97]. Black points are dijet MC (background) and $Z \to \tau\tau$ MC (signal).

Once the classifier is defined, working points are defined to select events according to a targeted efficiency for selecting τ_{had} candidates (signal efficiency). The loose, medium and tight working points are defined for 1p (3p) τ_{had} at an efficiency of 60%, 55% and 40% (50%, 40% and 30%) respectively. The performance is summarised in 4.4.

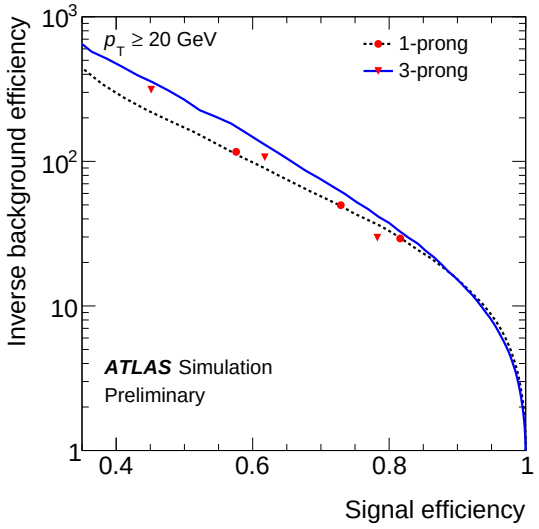


Figure 4.4: The receiver operating curve for the identification algorithm. This measures the signal efficiency against the background rejection when using a selection based on the BDT score. A larger area under the curve corresponds to a better discriminator [97].

Electron Discrimination

Electrons also have a very similar detector response as a one-pronged τ_{had}. The signature of a single track with a narrow cone of calorimeter deposits is very similar to that of the τ_{had}. As the identification algorithm is targeted at reduction of QCD jets, this is not effective against electrons and thus dedicated methods are required. Two sets of discrimination criteria are used to reject the presence of electron which can be reconstructed as τ_{had}.

Firstly, an overlap removal method is used [97]. A τ_{had} candidate is rejected if within $\Delta R < 0.4$ there exists a reconstructed electron candidate which is identified with the loosest working point of the likelihood identification criteria. Since the true electron selection efficiency for this working point is approximately 95% the rejection power when vetoing these events is very effective for identifying τ_{had}.

In addition to this overlap removal a further BDT was employed to separate τ_{had} and electron signatures. The BDT was trained in a similar manner to that used in reconstruction of τ_{had} with $\sqrt{s} = 8$ TeV data [98]. The performance is summarised in 4.5 for $\sqrt{s} = 8$ TeV data.

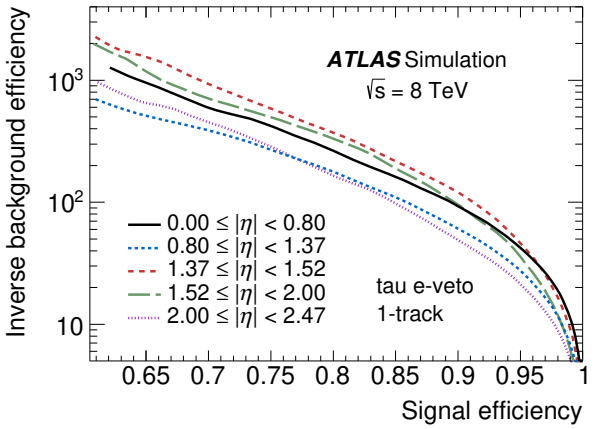


Figure 4.5: The receiver operating curve for the electron veto algorithm. This measures the signal efficiency against the background rejection when using a selection based on the BDT score. A larger area under the curve corresponds to a better discriminator [98]. Note this BDT was the algorithm used in $\sqrt{s} = 8$ TeV analyses.

4.3 Energy Calibration

The τ_{had}, in contrast to a QCD jet, contains a very distinct mixture of charged and neutral hadrons. Thus τ_{had} candidates benefit from a dedicated calibration to correct the energy which is initially from the jet reconstruction algorithm.

This calibration to the true τ energy, called the Tau Energy Scale or "TES", is achieved through the derivation of calibration constants, R. The calibrated energy at Tau Energy Scale is given by

$$E_{calib} = \frac{E_{reco} - E_{pileup}}{R(E_{reco} - E_{pileup}, \eta_{reco}, N_{prong})} \tag{4.1}$$

where E_{LC} is the energy of the jet seeding the reconstructed τ_{had} and E_{pileup} is a correction for pileup effects. R indicates the calibration constants which are measured as a function of E_{LC}-E_{pileup}, η and number of prongs of the τ_{had}.

The calibration curves obtained for τ_{had} from collisions at $\sqrt{s} = 13$ TeV are illustrated in figure 4.6.

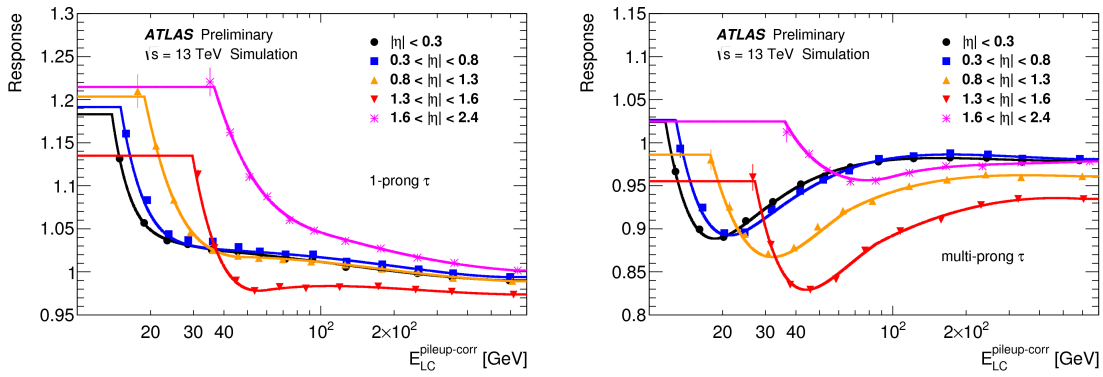


Figure 4.6: The calibration curves for the TES for one (left) and three (right) pronged τ_{had} [97].

These calibration constants are calculated based on MC estimates and in principle extrapolate the detector response to the true TES. In reality this is not the case as the TES observed in data can differ from the MC estimate after calibration. In section 4.4, a method of extracting the TES from data will be described and is the focus of the original work in this chapter.

The systematic uncertainty for the calibration to TES is taken from the relative shift of the ratio of the calibrated to true visible τ energy given a systematic variation. The uncertainties considered are: the non-closure of the method, the shower model, the response of the individual hadrons, underlying event, material effects, pileup and noise threshold effects.

The largest uncertainties which contribute to the total systematic uncertainty are the response of individual hadrons and the shower model. All other effects contribute at the $\sim 1\%$ level. The total systematic uncertainties are summarised in 4.7 and 4.8 for 1p and 3p τ_{had} respectively. The total variation is between 4-6%.

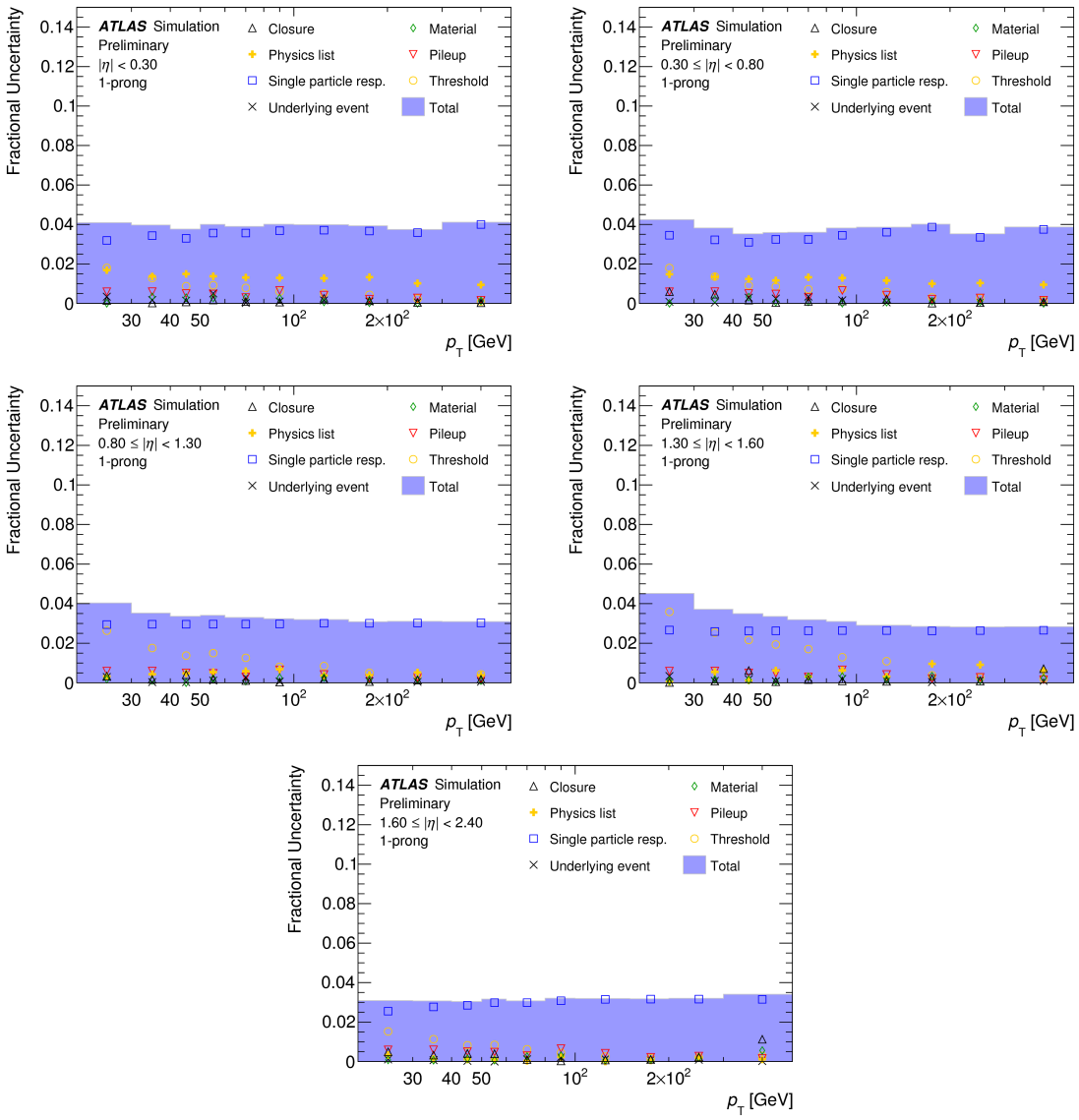


Figure 4.7: The sources of uncertainty for the calibration of the τ energy scale for 1p τ_{had} [97]. Each figure corresponds to a different region in $|\eta|$.

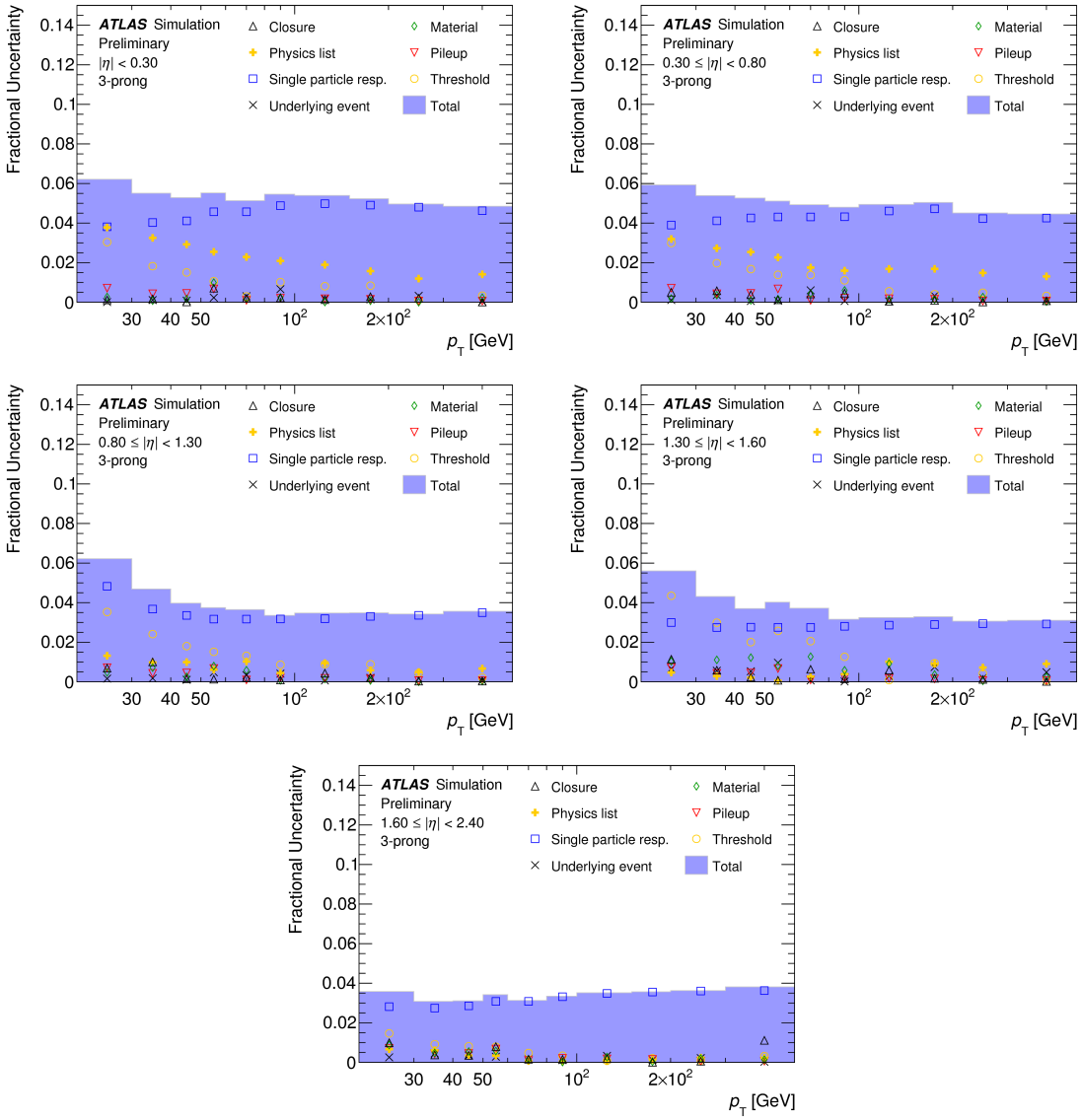


Figure 4.8: The sources of uncertainty for the calibration of the τ energy scale for 3p τ_{had} [97]. Each figure corresponds to a different region in $|\eta|$.

In addition to the simple baseline reweighting to produce the TES, a Boosted Regression Tree (BRT) is used for τ_{had} with $p_T < 250$ GeV. The BRT is a multivariate algorithm that attempts to regress the TES based on a number of input variables. This BRT combines information from a new reconstruction method[3] to minimise the energy resolution and the difference with respect to the generated p_T. This calibration is combined with the baseline TES calibration to maximise the effectiveness of the calibrations. Improving the τ_{had} resolution and accuracy of the τ_{had} reconstruction allows for the reduction in systematic uncertainties which have been shown to impact greatly the $H \to \tau\tau$ measurement [27].

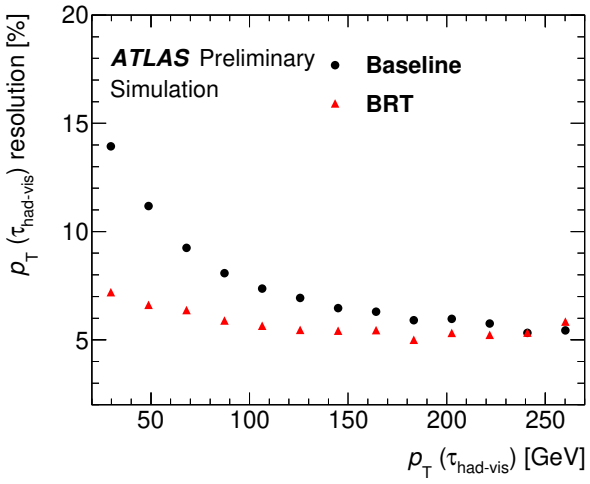


Figure 4.9: The p_T resolution of τ_{had} candidates with a given p_T for both baseline and BRT approaches [97].

[3]Further details are discussed in appendix A.

4.4 Data-driven Corrections

The calibrated τ_{had} have, thus far, only been corrected with respect to the MC expectation. Simulation of τ_{had} decays are particularly difficult as QCD at mid to low energies becomes incredibly difficult, hence semi-empirical parameterisations [99, 100] are used to model the τ_{had}. Often the energy, as well as identification efficiencies, ofτ_{had} observed in actual collisions will differ from that of these simulations. In order to have truly well calibrated τ_{had} candidates, the performance of each stage in the reconstruction is measured with respect to data, in particular for the TES. For this, a data-driven tag-and-probe technique is employed.

The tag-and-probe measurement detailed here attempts to select a high purity sample of $Z \rightarrow \tau_{\mu}\tau_{had}$ events in which one τ lepton decays to a muon and the other to hadrons (semileptonic decay). Since the reconstruction of muons very accurate and reliable, the "tagging" of an event using the identification of an isolated muon allows for an unbiased selection of events. The "probe" of the τ_{had} is then assessed to determine whether correction factors are required to bring the calibration in line with what is seen in data.

Measurements of data-driven corrections for τ_{had} identification, trigger, electron veto, efficiencies and TES via the tag-and-probe method share similar datasets, event selections and background estimation. The focus of this section will be development of a tag-and-probe method for the TES calibration in $\sqrt{s} = 13$ TeV collisions. Results of other $Z \rightarrow \tau_{\mu}\tau_{had}$ tag-and-probe methods are presented in section 4.4.3.

The dataset used for this measurement is based on $\sqrt{s} = 13$ TeV collisions collected in 2015 with total integrated luminosity of 3.2 fb^{-1}. The dataset is filtered by requiring that the muon pass the lowest p_T threshold single-muon trigger with minimum $p_T > 20$ GeV.

4.4.1 Event Selections

The event selections are intended to isolate the $Z \rightarrow \tau_{\mu}\tau_{had}$ events in order to concisely measure the TES shift as well as identification efficiencies. The selections target the rejection of backgrounds from multijet and W+jets events as well as events involving top quark production. The latter process can produce the μ and τ_{had} from the actual decay. Conversely, multijet and W+jets events produce the signature where either the τ_{had} is in fact a QCD jet or where the μ comes from a secondary b-quark decay.

The muon and τ_{had} candidates are reconstructed in the manner detailed in section 3.2.4. A minimum threshold of $p_T > 22$ GeV for the muon to ensure the muon p_T is on the trigger plateau as well as a gradient isolation working point is passed. Backgrounds are suppressed with a number of selections which are summarised in table 4.1.

- The gradient isolation requirement rejects events where the μ is from a non-prompt decay within a jet. This suppresses multijet events.
- The τ_{had} the is required to pass the *medium* working point with an additional requirement of BDT > 0.3 to reduce the presence of QCD jets.
- The muon and τ_{had} are required to have opposite electric charges[4]. This reduces the presence of QCD jets and small amount of top and W+jets backgrounds.
- W+jets events are suppressed primarily by requiring events to have a large transverse mass. The transverse mass is defined as $m_T = \sqrt{2p_T(\mu) \cdot E_T^{miss}(1 - \cos\Delta\phi(\mu, E_T^{miss}))}$ and forms a Jacobian peak [101] for W+jets events just below the W boson mass.
- $\Sigma\cos(\Delta\phi) = \cos\Delta\phi(\mu, E_T^{miss}) + \cos\Delta\phi(\tau_{had}, E_T^{miss})$ is used to reject W+jets and multijet events. This variable is sensitive to the direction of the E_T^{miss} with respect to the μ and τ_{had}. The majority of the signal is expected to peak about zero whereas the W+jets and multijet components have long tails extending past $\Sigma\cos\Delta\phi < 0$.

An additional selection on the mass window is warranted in the case of the efficiency measurements to purify the amount of $Z \to \tau_\mu\tau_{had}$. For the TES measurement no mass window constraint is required in order to fully measure the line shape of the Z mass.

Selection	Identification	Trigger Efficiency	TES
m_T	> 50 GeV		
$\Sigma\cos(\Delta\phi)$	> -0.1	> -0.5	> -0.5
m_{vis}	$\in [45, 80]$ GeV	$\in [45, 80]$ GeV	-
τ_{had} ID	-	various	Medium

Table 4.1: Event selections used for tag-and-probe measurements which derive data-driven corrections for the identification of, triggering with, and energy calibration for τ_{had}candidates. Note that for the measurement of identification efficiencies, this is performed for each of the τ_{had} working points.

[4]Events with candidates of the same charge are used for the background estimation.

4.4.2 Background Estimation

All backgrounds, with the exception of multijet, are simulated with MC summarised in the table 4.2. For all MC simulation, POWHEG [102–105] is used for the matrix element calculation and either PYTHIA6 [106] or PYTHIA8 [107] are used for the parton shower. The PDF set interfaced with the matrix element calculation and parton showering were the CT10 [108] and CTEQ6L1. In the case of V+jets an AZNLO tune is used for the underlying event whereas PERUGIA2012 [109] is utilised for the other samples.

The background estimation method utilises a data-driven technique to model events where the τ_{had} candidate is actually a QCD jet.

Process	Generator		PDF Set		Tune	Order
	ME	PS	ME	PS		
V+jets	POWHEG	PYTHIA8	CT10	CTEQ6L1	AZNLO	NNLO
$t\bar{t}$	POWHEG	PYTHIA6	CT10	CTEQ6L1	PERUGIA2012	NNLO+NNLL
Single top	POWHEG	PYTHIA6	CT10	CTEQ6L1	PERUGIA2012	NNLO

Table 4.2: Monte Carlo generators used for various stages of the simulation for the dominant signal and background processes considered. ME refers to the matrix element and PS refers to parton shower. The last column represents the order of the expansion in QCD or QED. The "tunes" column refers to the tuning of parameters required for the simulation of the parton showering.

The background estimation method detailed here apply for the measurement of TES shift as well as trigger efficiency. For the identification efficiency, the background estimation is a bit more involved as the proportion of backgrounds changes drastically with respect to the identification working points. More details for this are in [110].

The two major backgrounds, multijet and W+jets, to this analysis are due to QCD jets. The backgrounds and the $Z \to \tau_\mu \tau_{had}$ signal process are modelled with MC with the exception of the multijet background. The estimate of the multijet contribution is based on the so-called data-driven "OS-SS" method. This relies on the differences in correlation of the charges of the decay products between signal and background.

The signal of $Z \to \tau_\mu \tau_{had}$ almost entirely consists of events where the charges of the muon and τ_{had} are oppositely charged (OS). Thus a template in which the muon and τ_{had} candidates have the same charge (SS) is used to model the multijet background. This template is scaled by a normalisation factor, R_{QCD}, which corrects the normalisation to that expected by OS events. R_{QCD} is measured in multijet enriched regions defined with the same selections as the signal but where muons are required to fail the gradient isolation requirement. R_{QCD} is

defined as the ratio of events in the multijet enriched regions:

$$R_{QCD} = N_{OS}^{multijet\ CR}/N_{SS}^{multijet\ CR}, \tag{4.2}$$

where $N_i^{multijet\ CR}$ are the number of events in either the OS or SS mulitjet region.

To correct for double counting, the events from MC where the SS requirement is passed are subtracted from the total background. Separate correction factors, k_{OS} and k_{SS}, are derived in order to correct the normalisation of the W+jets background. The k factors are defined as the correction factor needed to scale the W+jets so that the total estimate matches data. The k factors are calculated in W+jets enriched control regions where m_T requirement is inverted ($m_T > 50$ GeV).

The overall estimate for the total number of events is expressed as:

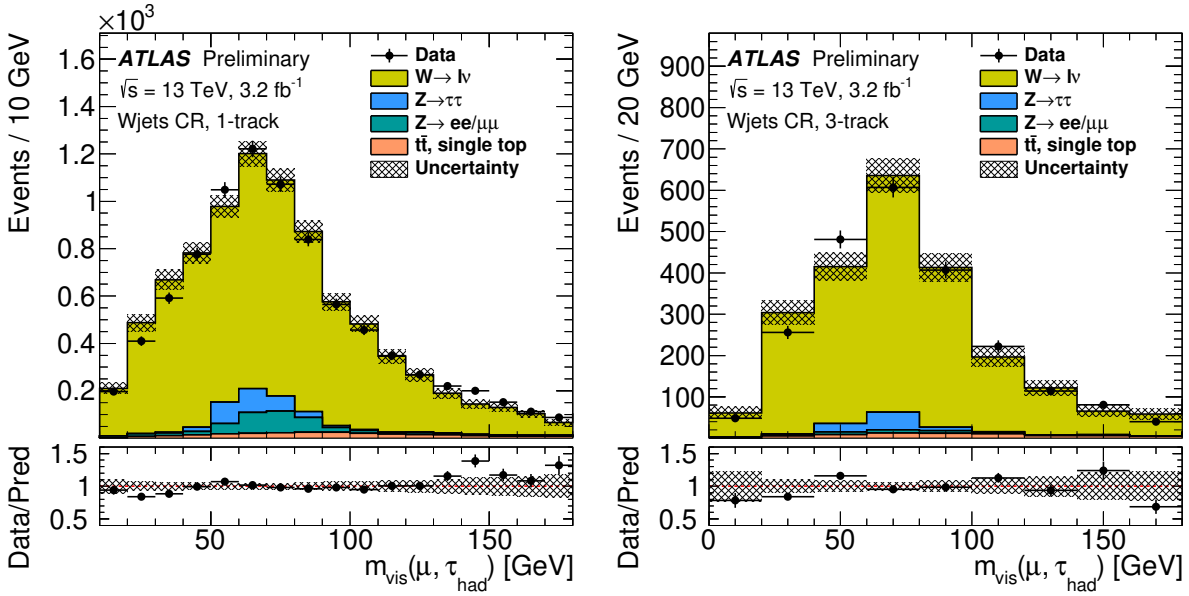
$$N_{OS}^{total} = R_{QCD} \cdot N_{SS}^{data} + N^{W+jets} + \Sigma N_{MC}^{process} \tag{4.3}$$

$$\text{where } N^{W+jets} = k_{OS} \cdot N_{MC,OS}^{W+jets} - k_{SS} \cdot R_{QCD} \cdot N_{MC,SS} \tag{4.4}$$

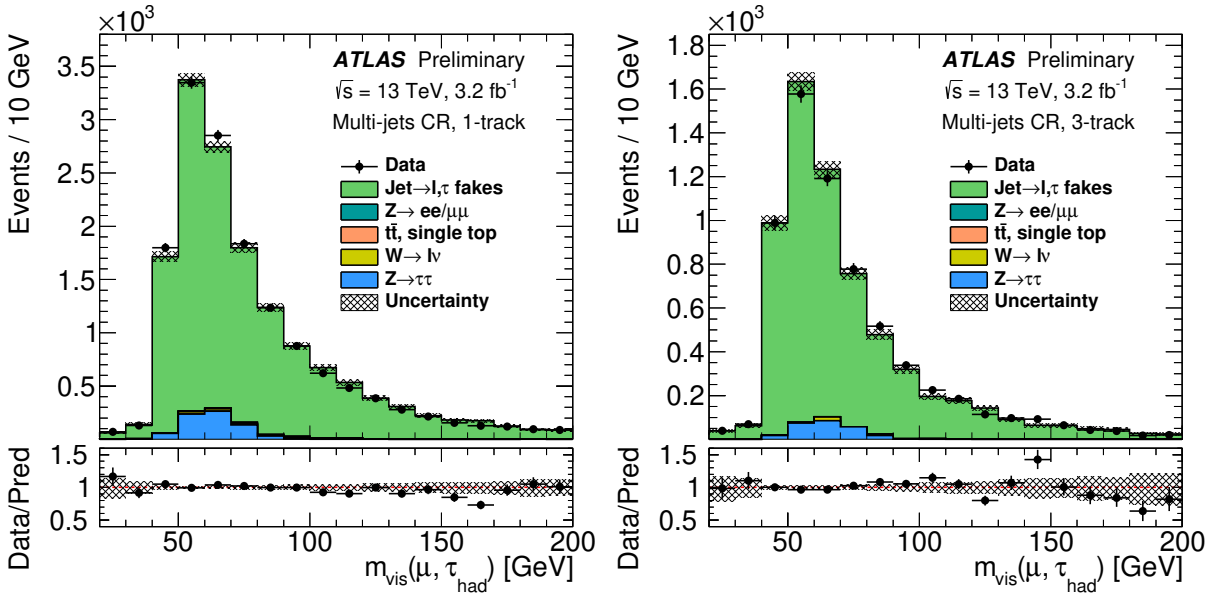
$$\text{and } N_{MC}^{process} = N_{MC,OS}^{process} - R_{QCD} \cdot N_{MC,SS}^{process} \tag{4.5}$$

and the summation runs over all other processes $Z \to \tau\tau$, $Z \to \ell\ell$ and top backgrounds. The data-driven factors are also derived separately for 1p and 3p events.

The quality of the modelling can be seen in figure 4.10, for two background enriched control regions. Good agreement is found in W+jets and multijet control regions for both events with 1p and 3p τ_{had}.



(a) W+jets Control Region



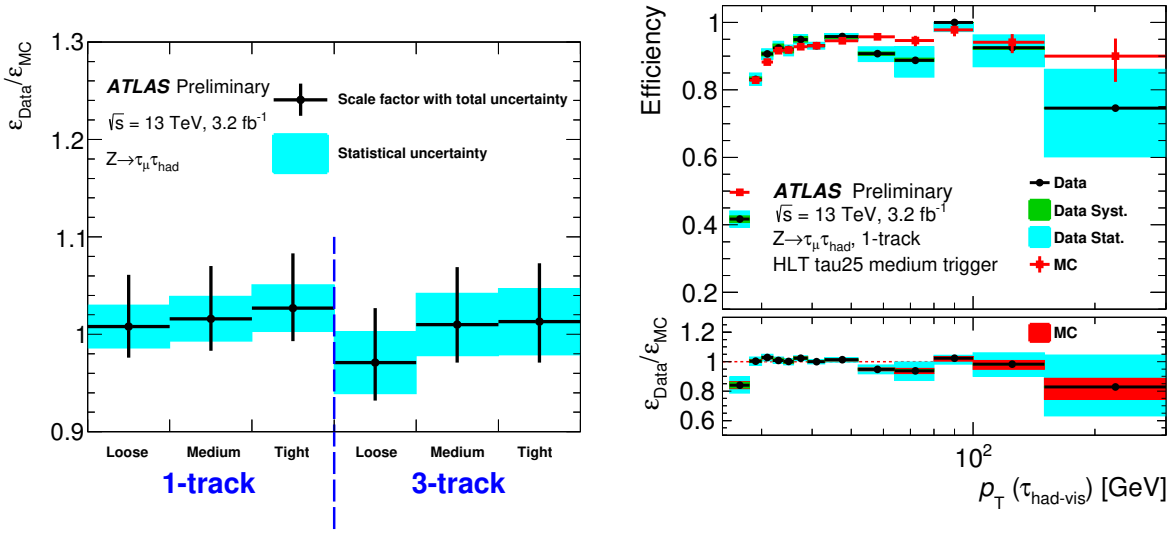
(b) Multijet Control Region

Figure 4.10: The m_{vis} distribution in various control regions. Plots left show events with 1p τ_{had} and plots right show those with 3p τ_{had}. The uncertainties are only statistical.

4.4.3 Efficiency Corrections

Corrections for the difference in efficiency is defined as $\epsilon_{data}/\epsilon_{MC}$, where ϵ_{data} is the efficiency in data and ϵ_{MC} is the efficiency in MC. For the identification efficiency, this ratio is calculated for passing a ID working point given a number of prongs. The efficiencies are measured using a fit to the number of track distribution before and after identification. The scale factor per identification working point and the τ_{had} track multiplicity is presented in figure 4.11a. The total systematic uncertainty is approximately 5% and is dominated by effects associated with the template modelling background estimation.

The trigger efficiency is measured as a function of $p_T(\tau_{had})$ and is calculated for the efficiencies of passing the L1 and HLT trigger. The results are presented in figure 4.11b. The dominant systematic uncertainties are due to the modelling of the jet fake background which are 3% for 1p and 7-8% for 3p τ_{had}. The statistical uncertainty is a secondary effect at ∼2% for 1p and 3-8% for 3p.



(a) Identification efficency

(b) Trigger Efficiency

Figure 4.11: Correction factors for identification efficiency (left) and trigger efficiency (right) [110]. Identification correction factors are depicted for 1p and 3p and various working points. Trigger efficiency (for 1p τ_{had}) is shown as a function of $p_T(\tau_{had})$ with the correction factor in the lower panel.

4.4.4 Tau Energy Scale Extraction

For the measurement of the TES correction, the line shape of the invariant mass of the $\mu - \tau_{had}$ pair, known as m_{vis} is used. This invariant mass peaks just below the Z mass of 90 GeV and is sensitive to the TES.

As the muon momentum is reconstructed to a very high precision, only a poorly calibrated TES can cause a mismodelling in the invariant mass peak. The effect of shifts in the TES results in a shift in the peak position of the invariant mass [98]. For deviations of less than 10% in the TES, this behaviour is linear. Thus, by measuring the difference in the invariant mass line shape between data and MC, one can measure the TES in data.

A shift parameter α is defined such that:

$$E_T^{data} = (1 + \alpha)E_T^{MC} \tag{4.6}$$

where E_T^{data} represents the TES observed in data and the E_T^{MC} is the TES as is expected from MC after the initial TES calibration described in equation 4.1. To measure the TES, samples are created with different values of α varying from -0.05 to 0.05. The assessment of the compatibility of a TES shift to data is done by calculating a χ^2 value:

$$\chi^2(\alpha, f) = \Sigma \frac{N_i^{data} - f N_i^{Z\to\tau\tau}(\alpha) - N_i^{bkg}}{N_i^{data} + f^2(\Delta N_i^{Z\to\tau\tau}(\alpha))^2 + (\Delta N_i^{bkg})^2}, \tag{4.7}$$

where the summation is performed over each bin of the m_{vis} distribution. N_i^{bkg} and ΔN_i^{bkg} are the number of events estimated for a process and the overall uncertainty respectively. The f parameter is introduced to account for normalisation difference in the $Z \to \tau\tau$. The measured value of the TES shift is given by the minima of all χ^2 values for any given normalisation f.

Systematic Uncertainties

Systematic uncertainties are evaluated by recalculating the χ^2 with a variation on affected background/signal samples and then remeasuring the minima. The difference with respect to the nominal value for the α is taken as a systematic uncertainty. The uncertainties are summarised in table 4.3.

Background modelling uncertainties are evaluated by varying the values of R_{QCD} and k_W up and down by their respective statistical uncertainty. Additionally tests of rebinning the width and varying the ID are also included in this. This is the leading systematic uncertainty.

Subleading effects are due to the τ_{had} energy resolution and muon reconstruction. Smaller uncertainties due to, for example E_T^{miss}, were found to be negligible.

Source	Uncertainty (%)	
	1p	3p
Background Modelling	1.1	2.9
Tau energy resolution	0.3	0.6
Muon reconstruction	0.2	0.6
Total	1.2	3.0

Table 4.3: The leading systematic uncertainties for the measurement of the insitu TES calibration. Values represent the difference with respect to the nominal measured α value. All systematic uncertainties smaller than 0.1% are not represented here.

Results

The distribution of the invariant visible mass with the corrected TES is presented in figure 4.12. It is evident that only a small shift in the TES is necessary to correct visible mass peak. Overall the measured TES correction is $\alpha = -0.7 \pm 0.8(\text{stat}) \pm 1.2(\text{syst})$ and $\alpha = -3.6 \pm 1.2(\text{stat}) \pm 3.0(\text{syst})$ for 1p and 3p respectively. For the case of calibration using the BRT the measured values are $\alpha = -0.95 \pm 0.9(\text{stat}) \pm 1.7(\text{syst})$ and $\alpha = -3.1 \pm 1.1(\text{stat}) \pm 1.6(\text{syst})$. This measurement ensures that τ_{had} energies are well calibrated. The uncertainties calculated here will be important in measurements involving τ_{had} decays.

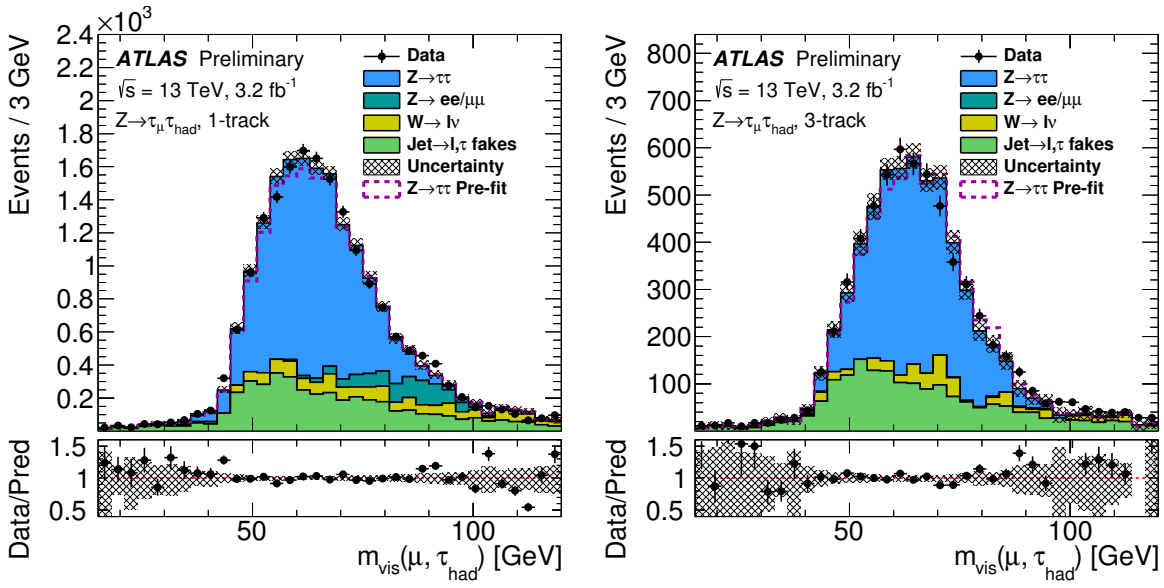


Figure 4.12: The m_{vis} distribution for 1p (left) and 3p (right) events. The purple distribution is the total estimate with the $Z \to \tau\tau$ uncorrected with respect to the measurement. The corrected $Z \to \tau\tau$ is displayed in solid blue. The uncertainties contain only the statistical uncertainty.

5

Measurement of the $H \to \tau\tau$ Cross-section

The measurement of the Higgs boson's τ coupling is crucial in assessing the mechanism for the generation of fermion masses. The latest measurement from ATLAS, prior to the results detailed here, was derived from $\sqrt{s} = 8$ TeV collisions [27]. Evidence for the existence of this coupling was established with a 4.5σ excess above the background-only hypothesis [27]. This was achieved utilising a multivariate classifier as the discriminating variable. This coupling was subsequently discovered with the CMS experiment once the $\sqrt{s} = 7, 8$ TeV datasets are combined with the $\sqrt{s} = 13$ TeV dataset with an integrated luminosity of 35.9 fb^{-1} [29]. The goal of this chapter is to outline the measurement which verified this coupling with the ATLAS experiment. This measurement analysed 36.1 fb^{-1} of $\sqrt{s} = 13$ TeV collision data.

The search is divided into three channels based on the decay mode of the two τ leptons, classified according to whether the τ leptons decay to leptons or hadrons. The semileptonic decay mode, with one lepton and one τ_{had} in the final state, contains the largest fraction of the $H \to \tau\tau$ decays (46%). This is compared to 12% and 42% for the fully leptonic and fully hadronic channels decays, respectively. The focus of this chapter is the measurement in semileptonic decays.

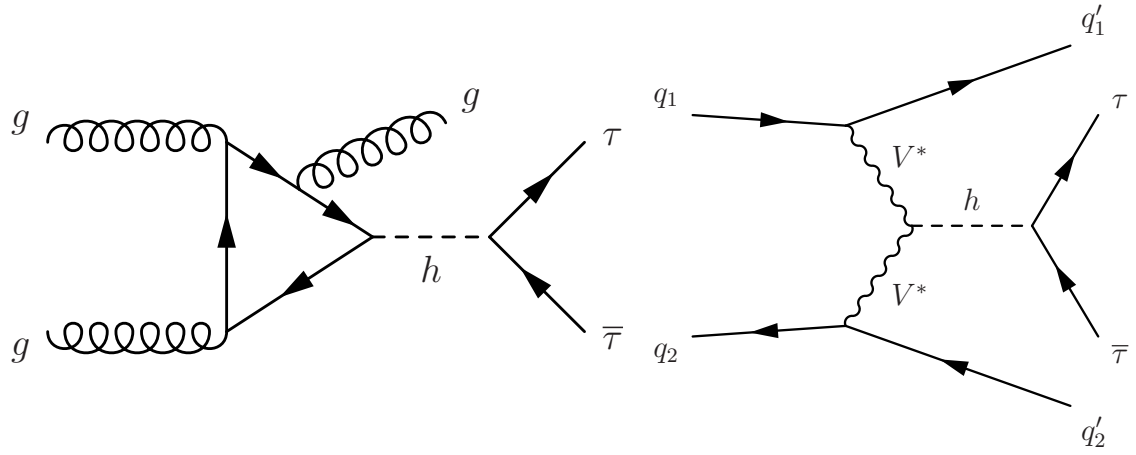
This channel is characterised by a large signal yield in the presence of challenging background processes. Thus the signal and background reconstruction will be addressed first before moving on the crux of the work presented: optimising the signal sensitivity and the signal extraction itself. The rejection of major backgrounds through targeted event selections will be detailed. Following this, there will be a description of the statistical model used to extract the outcome of this measurement.

5.1 Signal and Background

This section will discuss the experimental signature of the signal process and then move to describing the important background contributions. Note that details regarding the selection criteria used to define signal sensitive regions are described in section 5.5.

5.1.1 Signal Topology

As previously detailed in section 2.2, the Higgs boson is produced primarily via the ggH, VBF, VH and ttH modes at the LHC. This measurement targets only the ggH and VBF, the two leading production modes. These events constitute 90% of all produced Higgs bosons at the LHC. As will later be explained in section 5.5, events will be categorised based on the production mode: either a "boosted" ggH or VBF. The leading order Feynman diagrams relevant to this search are presented in figure 5.1.



(a) ggH + 1 jet (b) VBF

Figure 5.1: Leading order Feynman diagrams which are relevant to this measurement. Left is ggH production with 1 jet and right is the VBF production.

The Higgs boson produced via ggH is affected by large higher order corrections which are associated with the presence of jets through initial state radiation (ISR). In fact approximately half of all ggH produced Higgs bosons are associated with at least one jet. These jets, often featuring a large boost in the transverse plane, subsequently causes an equivalent transverse boost of the Higgs boson. The measurement presented here targets ggH events where the Higgs boson has a large transverse boost as many backgrounds tend to have a softer ditau transverse momenta. Figure 5.2 illustrates an example of such a ggH produced event.

While a significant proportion of ggH events are produced nearly at rest, such events are also typical of background processes. In particular, the large production rate of $Z \to \tau\tau$ events implies the gain in sensitivity with the inclusion of such events is minimal. Such a signal topology is excluded by requiring at least one high p_T jet to be reconstructed in the event.

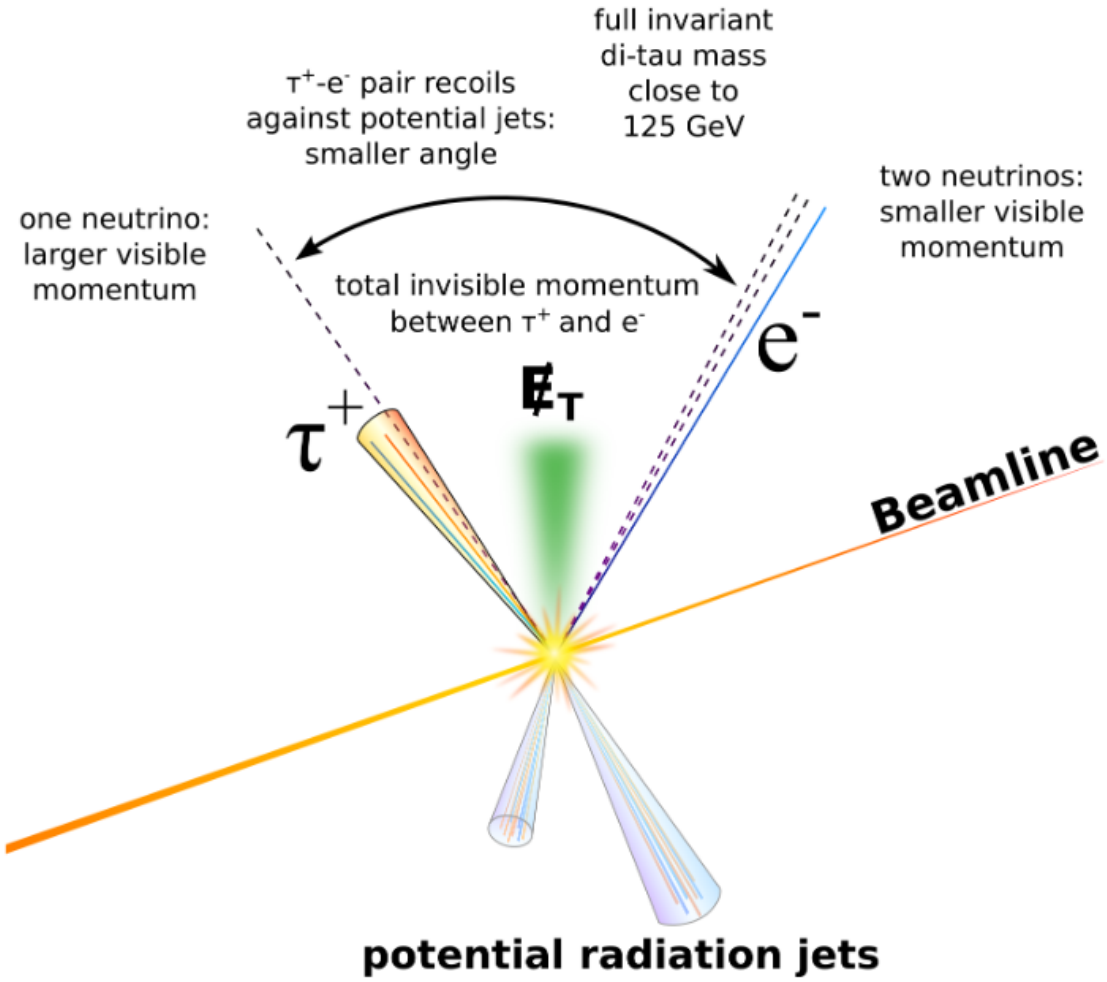


Figure 5.2: Schematic of a Higgs boson produced via ggH and decaying to a final state of a electron, τ_{had} and neutrinos

For the VBF topology, as there is no connection between the two sets of quarks, the resulting two jets tend to be collinear with respect to the initial momentum of the quarks. The resulting Higgs boson is produced from the fusion of the two vector bosons and will produce two τ leptons that decay in a pseudorapidity range between the two jets. The boost of the produced Higgs tends to be rather low in comparison to the ggH produced process however is still quite sizeable. In general, the characterisation of VBF topologies has a focus on the jet kinematics: particularly given a highly the presence of hadronic activity in the detector. This helps to reduce background contributions. A schematic is represented in figure 5.3.

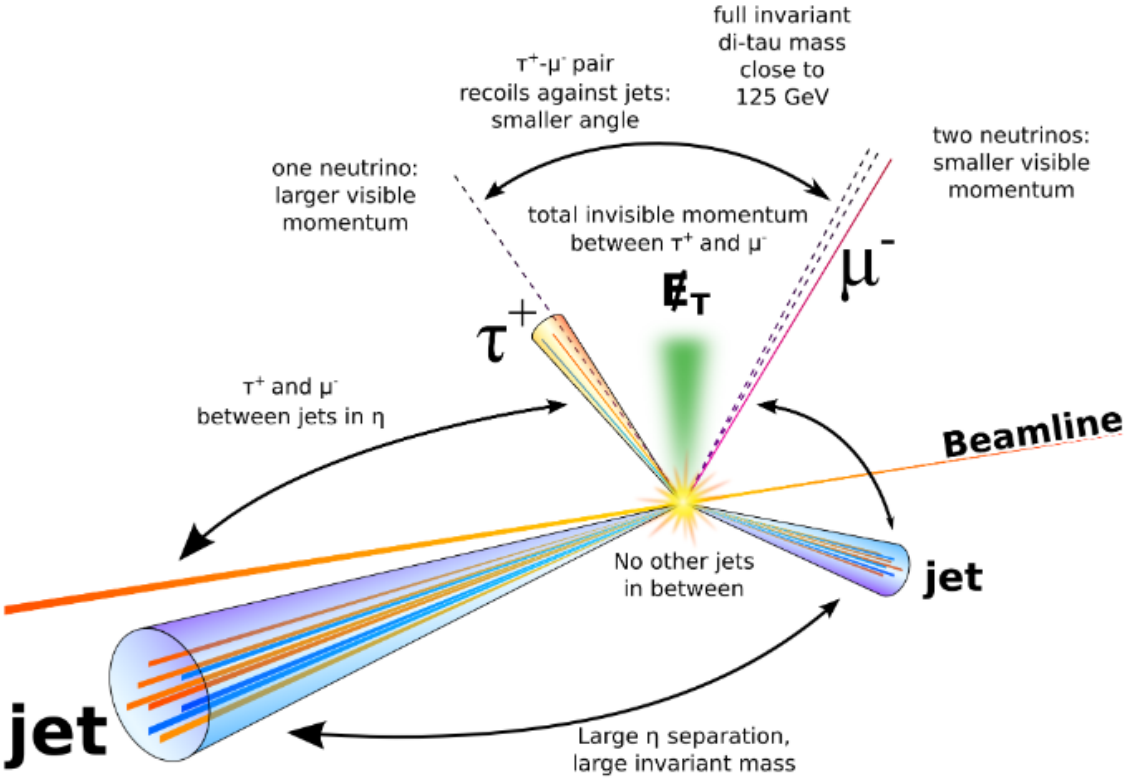
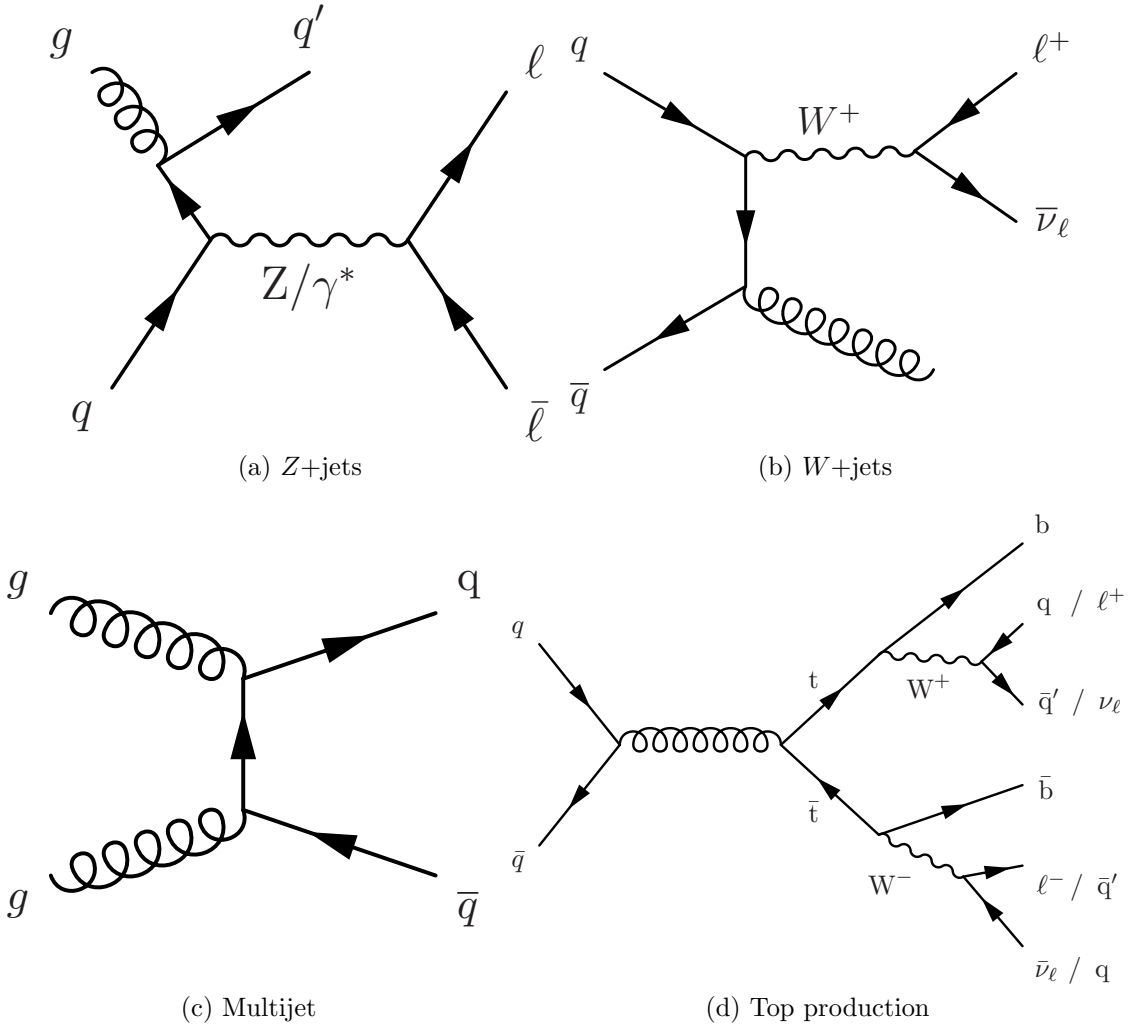


Figure 5.3: Schematic of a Higgs boson produced via VBF and decaying to a final state of a muon, τ_{had} and neutrinos

The semileptonic channel is a particularly useful one in measuring the $H \to \tau\tau$ coupling. With respect to the analysis of $\sqrt{s} = 8$ TeV collision data, the semileptonic and fully hadronic channels provided the best sensitivity [27]. This is expected to remain largely true with $\sqrt{s} = 13$ TeV collisions. The semileptonic channel benefits greatly from a lower probability of misidentification of final state particles compared to the fully hadronic channel and a higher trigger efficiency which exceeds 90% as compared to approximately 70% for the di-τ_{had} trigger [111].

5.1.2 Backgrounds

The experimental signature described in the previous section 5.1.1 can also be replicated by several background processes. These processes are categorised into either reducible and irreducible background sources. A summary of some Feynman diagrams representing background contributions are presented in figure 5.4.



(a) Z+jets (b) W+jets

(c) Multijet (d) Top production

Figure 5.4: Examples of Feynman diagrams which contribute to background sources. Here ℓ can represent either any lepton, including τ lepton.

Reducible sources of background are characterised by a final state where one or more particles is misidentified. The bulk of these background originate from W production in association with initial state radiation of one or more jets, or non-resonant QCD processes which produce multijet signatures. These are typically referred to as "jet fakes" as the τ_{had} candidate is a reconstructed jet which is misidentified as a τ_{had} candidate. In the case of multijet events, the lepton is misreconstructed from decay of a hadron within a jet. Primarily, the background rejection here is achieved through strict identification requirements on the τ_{had} candidate.

Smaller background sources originate from cases where a process can produce the same signature as the $H \to \tau\tau$ signal but also produces additional particles. These additional particles often include soft jets as well as lepton candidates which don't pass the reconstruction requirements. Included in these processes are the production of single top or top quark

pairs and the production of combinations of two electroweak bosons, henceforth referred to as "diboson production". Production of $Z \to ll$ + jets can also produce a similar signature. These are reduced through kinematic selections as the distribution of leptons and jets are typically quite different from the $H \to \tau\tau$ signature.

The more difficult background sources are those which are irreducible in nature. As all the particles are properly reconstructed, irreducible backgrounds only differ from the signal in the kinematics of the final state particles. The only background of this type is due to $Z \to \tau\tau$ decays. The main handles which reduce the contribution of this background are the reconstructed mass of ditau pair and selections which target the boost of the Higgs boson/Z boson. The latter was previously discussed in section 5.1.1. This reconstructed mass, the main discriminator for this measurement, is able to reliably resolve the Z and Higgs boson peaks to a reasonable degree. This will be discussed later in section 5.7.

Having discussed the experimental signature of the signal and the backgrounds which contaminate them, an analysis strategy is formulated to measure the $H \to \tau\tau$ cross-section.

5.2 Analysis Strategy

In order to measure the $H \to \tau\tau$ cross-section, a number of steps are required. The structure of the remainder of this chapter follows this procedure. Firstly a dataset must be collected in order to test the hypothesis (section 5.3). Following this, particle candidates (section 5.4) are selected. The modelling of signal and background processes (section 5.6) is then implemented. Before extracting the signal, a strategy for achieving optimal signal sensitivity must be devised (section 5.5). Systematic uncertainties are then evaluated 5.8 and the signal is extracted (section 5.9). The results are then presented in section 5.10.

The key to a well measured result lies in three parts: the background estimation, an optimised strategy for signal sensitivity and a robust signal extraction methodology. The background estimation is based on a method developed for analysis of the $\sqrt{s} = 8$ TeV dataset by ATLAS [27]. As this is not the focus of the author's work, only a brief overview is offered.

As the $H \to \tau\tau$ cross-section is expected to be dwarfed by the backgrounds, good discrimination of backgrounds is of paramount importance. Despite the multivariate approach (used in [27]) potentially outperforming a traditional cut-based approach, this measurement is aimed at bridging the gap to a differential cross-section measurement. As such, extracting signal from a classifier score will cause complications when attempting to make a fiducial or differential cross-section measurements. Instead this measurement's optimisation of signal sensitivity is achieved through a cut-based analysis which has been greatly optimised with respect to [27]. This optimisation strategy involves targeted event selections which exploit unique signal topologies to actively reject background contributions.

The signal is then extracted from a maximum likelihood fit which is binned according to the key discriminator in this measurement, a reconstructed Higgs boson mass. The estimation of this mass is described in section 5.7. The extraction of the signal involves a complex statistical model involving over 200 parameters.

5.3 Dataset and MC Simulation

The data used in this measurement was collected by the ATLAS detector during proton-proton collisions between 2015 and 2016. The centre-of-mass energy of these collisions was $\sqrt{s} = 13$ TeV and the total integrated luminosity is 36.1 ± 1.0 fb^{-1}.

The data which was used for the semileptonic channel was collected utilising the single muon and single electron triggers. These triggers impose a requirement on the minimum lepton p_T as well as isolation restrictions. A summary of these requirements is presented in table 5.1.

For the 2015 data-taking, the online thresholds on the reconstructed lepton p_T were 24 and 20 GeV for electron and muon triggers, respectively. As a result of increasing instantaneous luminosity during 2016 data taking, a higher p_T threshold of 26 GeV as well as additional isolation requirements were imposed for events to pass the electron and muon triggers. Rather loose isolation selections with high efficiency ($> 95\%$) [82, 85] are made to further reduce the rate of data taking.

To ensure the highest trigger efficiency, the offline thresholds on the lepton candidates for the 2015 (2016) dataset are 25 (27) and 21 (27.3) GeV for electron and muon triggers, respectively. Similarly for the isolation requirement, a gradient isolation criteria [82, 85] is used to ensure events are selected only with a tighter requirement than specified for the trigger. The efficiency of this working point varies in p_T between 90% and 99% at 25 and 60 GeV, respectively.

Trigger	p_T threshold $\langle$GeV$\rangle$		Isolation	
	2015	2016	2015	2016
Single Electron	25 (24)	27.3 (26)	Gradient (None)	Gradient (Loose)
Single Muon	21 (20)	27 (26)	Gradient (Loose)	Gradient (Medium)

Table 5.1: Momentum thresholds and isolation for datasets collected with the electron and muon triggers. In brackets are the online thresholds.

MC simulation is used for all of signal processes and most of the important background contributions. The matrix elements for the major signal processes are generated with the Powheg-Box v2 generator [102–105]. For ggH signal, this was calculated to NNLO accuracy using NNLOPS [112].[1] VBF and VH signals are calculated to NLO accuracy with the MiNLO approach [113]. The ttH signal was simulated with MG5_aMC@NLO v2.2.2 [114] to NLO accuracy in QCD and later renormalised to be NLO The total cross-section of the ggH processes was renormalised with a calculation at N^{3}LO (next-to-next-to-next-to-leading

[1] The accuracy is initially NLO but is reweighted to NNLO as a function of the Higgs boson rapidity.

Process	Monte Carlo generator	PDF	UEPS	Cross-section order
ggH	POWHEG-BOX v2	PDF4LHC15 NNLO	PYTHIA 8.212	N^{3}LO QCD + NLO EW
VBF	POWHEG-BOX v2	PDF4LHC15 NLO	PYTHIA 8.212	∼NNLO QCD + NLO EW
VH	POWHEG-BOX v2	PDF4LHC15 NLO	PYTHIA 8.212	NNLO QCD + NLO EW
ttH	MG5_aMC@NLO v2.2.2	NNPDF30LO	PYTHIA 8.1	NLO QCD + NLO EW
V+jets	SHERPA 2.2.1	NNPDF30NNLO	SHERPA 2.2.1	NNLO
Diboson	SHERPA 2.2.1	NNPDF30NNLO	SHERPA 2.2.1	NLO
$t\bar{t}$	POWHEG-BOX v2	CT10	PYTHIA 6.428	NNLO+NNLL
Single top	POWHEG-BOX v1	CT10F4	PYTHIA 6.428	NLO

Table 5.2: Monte Carlo generators used to describe all background and signal processes. The configuration of PDF set parton showering model, hadronisation and underlying event (UEPS) are detailed here. The last column describes order of the total cross-section calculation is given.

order) in QCD [115–118] whereas VBF was renormalised to approximately NNLO [119–121] and VH to NNLO in QCD [122–124]. Electroweak corrections of NLO accuracy are applied to all signal samples. The PDF4LHC15 is utilised as the PDF set in the calculation for all but the ttH signal. The PDF4LHC15 is tuned with AZNLO [125] parameters and the CTEQ6L1 PDF set [126]. For ttH signal, the NNPDF30LO [127] PDF set interfaced with the matrix element calculation. PYTHIA 8.212 [107] was used in the simulation of the parton showering for all signal processes.

For production of W and Z bosons, either a single W/Z boson or in pairs, is generated with the SHERPA 2.2.1 generator which merges the matrix-element calculation with the parton shower model. The matrix elements are calculated with the Comix generator at LO [128] and OpenLoops generator at next-to-leading order (NLO) [129]. For the single W/Z production with a number of jets, the matrix elements are calculated to NLO for up to two partons and up to four at leading order (LO) with the overall cross-section taken from an next-to-next-to-leading order (NNLO) prediction from FEWZ [130, 131]. The NNPDF30NNLO [127] PDF set is used in both sets of samples.

The generation of the $t\bar{t}$ background, is modelled with POWHEG-BOX v2 interfaced with PYTHIA 6.428 [106] for the parton shower. For the single top quark background is generated using POWHEG-BOX v1 [132, 133], again with PYTHIA 6.428 for the parton shower. CT10 [108], which incorporates a 4-flavour quark scheme, and CTEQ6L1 [134] PDF sets are used for the calculation of matrix element and parton shower respectively. The tune used for the parton showering is PERUGIA2012 [109]. The cross-section for the $t\bar{t}$ process is calculated to NNLO in QCD with soft-gluon resummation to the level of next-to-next-to-leading log (NNLL) with TOP++2.0 [135].

5.4 Particle Candidate Selections

The signal topology consists of one electron or one muon, a τ_{had} candidate as well as missing energy. The event must also contain one or more reconstructed high p_T jet. A particle candidate selection is required for all of electrons, muons, τ_{had}, jets and missing transverse energy. A brief summary is offered here, highlighting differences with respect to the more general description in section 3.2.4.

One electron or muon as well as a τ_{had} candidate are required. Lepton candidates are required to have a p_T threshold and isolation requirements in accordance with the constraints from the trigger (see 5.1). For both muons and electrons the identification requirements from section 3.2.4 are tightened to require *medium* identification. For electrons the efficiency of the medium working point varies from 87% at E_T above 20 GeV to 95% at $E_T>$ 60 GeV and for muons this exceeds 96%.

The τ_{had} candidate with $p_T > 20$ GeV is required to pass a *medium* BDT working point to suppress the large contamination from jets which would otherwise pass the selections. The medium working point is a balance between high signal efficiency - 55% and 40% for 1p and 3p τ_{had} respectively - and good background rejection. A selection on an electron-τ_{had} discriminator is required to reject electrons which are misreconstructed as τ_{had}. The selection efficiency for one pronged τ_{had} candidates is 95%.

A jet is selected if it has $p_T > 20$ GeV within a $|\eta| < 4.9$ region. A vertex tagging algorithm is used to reject candidates consistent with pileup events. A b-tagged jet is considered if it passes a $p_T > 25$ GeV selection and is within the acceptance of $|\eta| < 2.4$. The b-jets are identified utilising a multivariate algorithm which is detailed in section 3.2.4: the efficiency of selecting such jets is 85%.

Missing transverse energy is present due to neutrinos in the final state which escape detection by ATLAS. The E_T^{miss} is reconstructed by summing the p_T of particle candidates as specified in section 3.2.4. This includes the addition of "soft terms" originating from clusters and tracks not associated with any other particle candidates.

The signature of different particle candidates may look very similar to other reconstructed candidates. For example, a τ_{had} can be reconstructed as both a τ_{had} and jet candidate. In the case where a particle can be reconstructed as multiple types of candidates, a methodology known as "overlap removal" is utilised to uniquely and consistently define what the type of particle is. If a particle can be reconstructed as multiple different types of particle candidate within $\Delta R < 0.2$ ($\Delta R < 0.4$ for jets), a fixed hierarchy dictates what the particle is. As muons are the most distinctive signature, the candidate is always chosen to be a muon if it can be reconstructed as such. Following this, the priority of signatures are electrons then subsequently τ_{had} and finally other jets.

5.5 Event Selections

This section will outline the approach taken to improve upon the ATLAS analysis of the $\sqrt{s} = 8$ TeV dataset through a well considered cut-based selection. It will also outline the selections used to determine background enriched regions, known as control regions (CR), used in background estimation and/or validation. The other channels will only enter when summarising the fully combined results in section 5.10.[2] All selections for the semileptonic channel are detailed in table 5.4.

Only events passing a loose "preselection" will be considered. This preselection is defined to broadly categorise events which are signal-like. This region aims to have a high signal acceptance but low purity. Selections are made to reduce the largely reducible backgrounds. More specialised selections are detailed in the following section to reject the larger backgrounds which remain in the preselection region. The selections for preselection are:

- A semileptonic signature of one lepton and one τ_{had} candidate are required with opposing charges.
- A selection on the transverse mass (m_T)[3] distribution is applied to reject W+jets events.
- A b-jet veto ensures events from top decays will be reduced.
- Angular separations (ΔR and $\Delta\eta$) between lepton and τ_{had} are required to reject multijet backgrounds.
- Additional selections on the E_T^{miss} and visible fraction of the τ momentum, x_i which is defined in 5.4, ensures that the missing energy is distributed as expected for signal.

Following the passing of preselection events are then categorised to focus on the specific production processes and to signal sensitive regions (signal regions - SR) to maximise the effectiveness of the signal extraction.

5.5.1 Signal Regions

Two categories of events are defined to target either the ggH or VBF signatures. A "boosted" category has a focus on ggH events which are produced in association with one jet whereas the "VBF" category targets the two jet VBF topology.

[2] More detail can be found in [3].
[3] See section 4.4.1 for details.

Inclusive Categorisation

The more distinctive production process is the VBF production due to its topology of two back-to-back jets in the forward region of the detector. The inclusive "VBF" category is then defined based on the jet topology. The requirements for an event to be included in the VBF category are:

- Two jets which are with a large $\Delta\eta$ separation.
- The two jets being in opposite in hemispheres (pointing in different directions along the beampipe).
- The Higgs boson system of the lepton and τ_{had}, being central between the two jets is required.
- A large mass calculated from the two forward jets ("dijet mass" - m_{jj}).

Following the categorisation of VBF-like events, ggH-like events with a large transverse boost from a recoiling jet are targeted in the "boosted" category. Events failing the VBF categorisation define the "boosted category". To emphasise the events where the Higgs boson is boosted, a selection is made on the transverse momenta of the Higgs boson candidate (${p_T}^H$), which is defined by the sum of four-vectors of the lepton, τ_{had} and E_T^{miss}. As previously discussed, this works to reduce the ratio of signal to $Z \to \tau\tau$ events by a considerable amount. In spite of this, this category remains largely dominated by $Z \to \tau\tau$ events.

Exclusive Signal Regions

The categorisation distinguishes events according to signal topology based on selections which were used to analyse the $\sqrt{s} = 8$ TeV dataset [27]. This result was derived primarily with a multivariate approach in order to maximise the discovery potential. However, as it is clear we have strong evidence already, the aim is now shifted away from discovery and more to a measurement of cross-section. As such, having clearly defined fiducial regions is more beneficial to such a task and as such a cut-based approach is primarily used for this measurement. The challenge is to retain the background suppression while retaining a high signal acceptance as defining a single signal region will often result in either low background suppression or low signal acceptance.

As such, one development since then has been to further develop multiple signal regions within these inclusive Boosted and VBF categories. More signal sensitive regions, simply "signal regions", are defined by selections targeting the rejection of the major background in the inclusive categories. In order to retain signal acceptance however, the events passing the inclusive categories but failing the more stringent requirements of the signal sensitive region are collected in another region. This allows for an effective selection scheme which does not reject any signal events.

VBF regions: To define a more signal sensitive region, the "VBF tight" region, a boosted Higgs boson candidate is required. This was shown in the fully hadronic analysis in the $\sqrt{s} = 8$ TeV analysis to demonstrate good discrimination with respect to backgrounds [27]. In addition a selection on the dijet mass is required to reject misidentified τ_{had} backgrounds. Remaining events compose the "VBF loose" region.

Boosted category: The "high p_T" region is defined requiring a high boost of the Higgs boson candidate and a small ΔR separation between the lepton and the hadronic tau. This selection rejects the non-resonant misidentified τ_{had} background. The "low p_T" region contains the remaining events.

A breakdown of the proportions of signal and background in the semileptonic signal and control regions are depicted in figure 5.5. Exact yields tabulated in table 5.3. It is evident that the more signal sensitive regions, boosted high p_T and VBF tight, contain a larger fraction of $Z \to \tau\tau$ background and lower amounts of misidentified τ_{had} backgrounds. These selections effectively reject the non-resonant misidentified τ_{had} background. The reconstructed mass ultimately proves critical in discrimination against $Z \to \tau\tau$.

Sample	Preselection	Boost	Boost High	Boost Low	VBF	VBF Tight	VBF Loose
Fakes	5290±40	2491±24	609±13	1882±21	172.0±6.7	86.1±5.2	85.8±4.2
$Z \to \tau\tau$	15660±120	9427±81	5314±55	4113±59	552±19	345±13	207±14
Top	295±10	194.0±8.3	59.1±4.8	134.9±6.8	25.0±3.2	18.0±2.7	7.0±1.7
Diboson	347.7±5.7	253.4±4.7	134.4±3.2	119.0±3.4	14.0±1.6	10.6±1.5	3.36±0.68
$Z \to ll$ + jets	760±110	273±36	108±15	165±33	21.2±4.7	11.1±3.9	10.1±2.6
W+jets	133±22	73±12	26.7±5.6	46±11	24±14	13.3±9.6	11.0±9.8
ggH	213.4±1.9	125.6±1.4	64.31±0.95	61.24±0.97	10.88±0.41	6.99±0.31	3.88±0.26
VBFH	69.72±0.39	25.92±0.24	13.83±0.17	12.09±0.16	32.48±0.27	25.36±0.24	7.12±0.13
VH	18.16 ± 0.36	13.27 ± 0.30	7.67 ± 0.22	5.60 ± 0.21	0.29 ± 0.05	0.16 ± 0.03	0.13 ± 0.04
ttH	0.75 ± 0.13	0.63 ± 0.12	0.20 ± 0.06	0.43 ± 0.10	0.05 ± 0.05	0.05 ± 0.05	0.00 ± 0.01
Background	22480±170	12712±93	6251±59	6461±72	809±25	484±17	325±18
Signal	302.1±1.9	165.36±1.4	86.00±0.99	79±1.0	43.70±0.49	32.56±0.40	11.14±0.29
Data	22836	12903	6347	6556	814	496	318

Table 5.3: Expected number of background events at various stages of event selection in the semileptonic channel. Only statistical uncertainties are listed.

	$e\tau_{had}$	$\mu\tau_{had}$
Preselection	Exactly one electron candidate No muon candidates BDT$>$ 0.15(0.05) for 1p (3p) τ_{had}	Exactly one muon candidate No electron candidates -
	Lepton and τ_{had} are oppositely charged $m_T < 70$ GeV No b-jet candidates Leading jet $p_T > 40$ GeV $E_T^{miss}>$ 20 GeV $\Delta R(l, \tau_{had}) < 2.5$ $\lvert\Delta\eta(l, \tau_{had})\rvert < 1.5$ $x_1 \in (0.1, 1.4)$ $x_2 \in (0.1, 1.2)$	
VBF inclusive	Pass Preselection Sub-leading jet $p_T > 30$ GeV $m_{jj} > 400$ GeV $\lvert\Delta\eta(j_0, j_1)\rvert > 3$ $\eta_{j_0}\dot{\eta}_{j_1} < 0$ Lepton and τ_{had} centrality between jets	
VBF signal regions	Tight	Loose
	Pass VBF inclusive $p_T^H > 100$ GeV $m_{jj} > 500$ GeV	Pass VBF inclusive Fail VBF Tight
Boosted inclusive	Pass Preselection Fail VBF inclusive $p_T^H > 100$ GeV	
Boosted signal regions	High p_T	Low p_T
	Boosted inclusive $p_T^H > 140$ GeV $\Delta R(l, \tau_{had}) < 1.5$	Boosted inclusive Fail High p_T

Table 5.4: Summary of selections defining the preselection, inclusive categories and signal regions the semileptonic channel. Note that "BDT" refers to the BDT to discriminate electrons from τ_{had}.

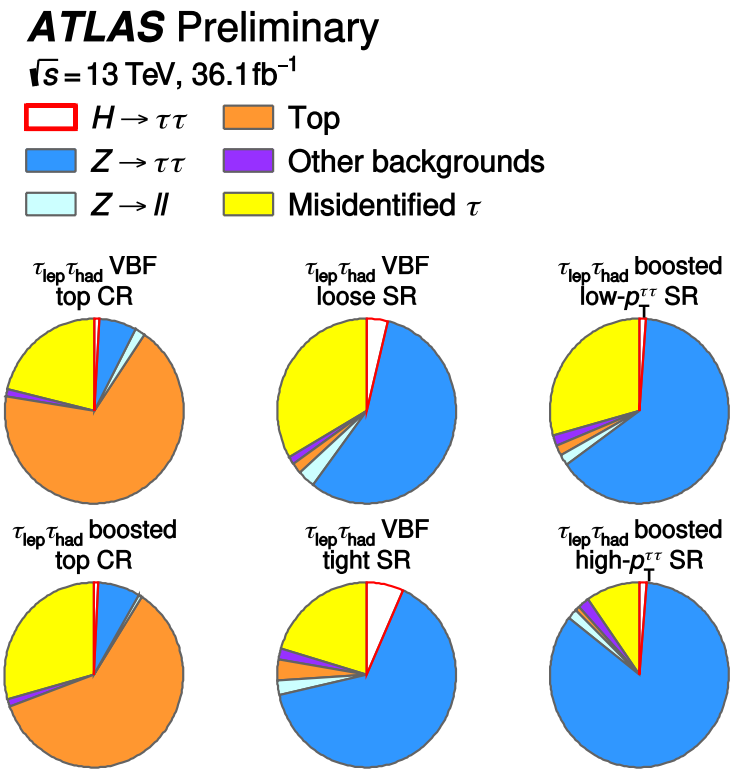


Figure 5.5: Expected proportions of signal and background components in both SR and CR of the semileptonic channel.

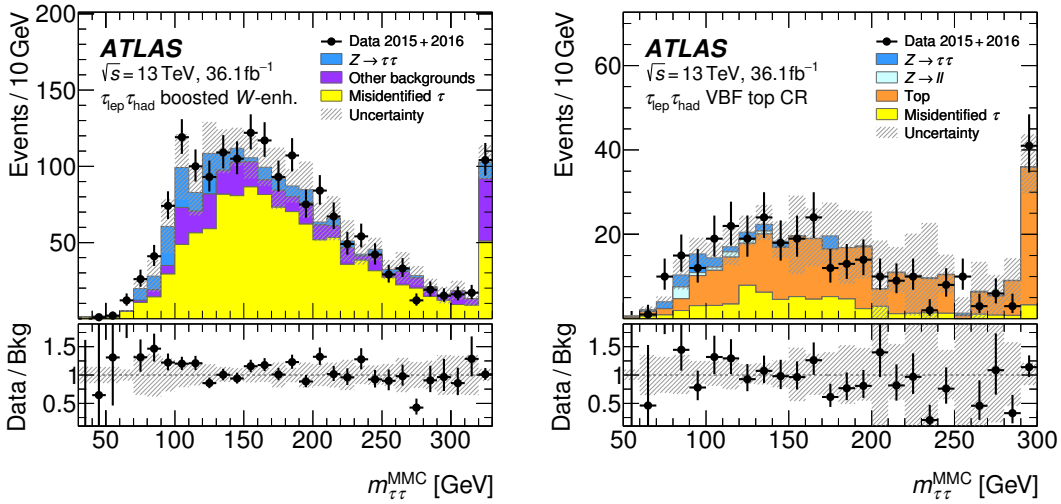
5.5.2 Background Control Regions

The modelling of backgrounds, either reducible or irreducible, requires the study of regions enriched in a particular backgrounds. These regions are used either in the background estimation or to check the modelling. These background enriched regions are defined by inverting signal selections such that a background has high purity with a low signal purity. For the relevant backgrounds, two such regions are defined: one for boosted and one for VBF. Regions are based on inverting selections on the inclusive boosted and VBF regions. The modelling of the $Z \to \tau\tau$ background is of particular importance and will be discussed in depth in section 5.6.3.

The presence of considerable top backgrounds, up to 4% of the background, calls for an enriched region to check the modelling of these backgrounds and correct for differences in normalisation. Top control regions are defined by inverting the veto on the presence of b-jets and lowering the m_T threshold to 50 GeV in order to increase the purity of top backgrounds. One example is illustrated in 5.6b showing good agreement between data and the top background estimate.

The estimation of backgrounds where a jet is misidentified as a τ_{had} is data-driven, and requires control regions for W+jets or multijets processes to be defined. W+jets control regions are defined by inverting the transverse mass requirement, allowing events around the W mass peak. Multijet control regions are defined by inverting the lepton isolation. These will be discussed later in section 5.6.1.

The W+jets control regions are also used order to check the modelling of the misidentified τ_{had} background. As evident in figure 5.6a, the data-driven background estimation models well the observed data. This builds confidence that the background estimation has been performed correctly and will demonstrate a similar degree of agreement in the SR.



(a) W+jets enriched region (Boosted) (b) Top enriched region (VBF)

Figure 5.6: Example distributions of the m_{MMC} variable in the W+jets and top enriched control regions following the extraction of normalisations for the top background. The lower panel is the ratio between data and background with the uncertainty band containing all systematic and statistical uncertainties.

5.6 Background Estimation

Background contributions must be correctly estimated in the signal regions before signal extraction can be performed. In most cases, MC simulation is insufficient to describe the backgrounds. Data-driven methods are preferred for estimating backgrounds which have a large presence in the signal regions. This section will describe how these contributions are estimated, validated and how systematic uncertainties associated to the modelling are addressed.

5.6.1 Jets misidentified as τ_{had} Background

A large background from events where a jet is misreconstructed as a τ_{had} exists. This is due to a large cross-section for multijet and W+jets production and the substantial background acceptance from jets misreconstructed as τ_{had}. These backgrounds are labelled misidentified τ_{had} or "fake" in this chapter. Note that it was found that leptons misidentified as τ_{had} contributed, at most, 5% to the total yield and thus the modelling of this contribution was based on MC estimates.

The data-driven method used to model the backgrounds producing misidentified τ_{had} is known as the "fake factor" method. For a given signal region, the fake factor method models the distribution of a variable by using events in a region enriched in jets. The distribution, which will be termed a template, is corrected to match the normalisation expected in a given signal region by applying a transfer factor. This procedure can be briefly described in the following equation:

$$N_{\text{expected}} = N_{\text{anti-}\tau_{had}} \times FF, \tag{5.1}$$

where N_{expected} is the number of events expected in a given region, $N_{\text{anti-}\tau_{had}}$ is the number of events in the template region and FF encodes the transfer factor which is binned in a number of variables. The terms will be explained below.

The region where the template is derived is a jet enriched region whereby "anti-τ_{had}" candidates are required rather than τ_{had}. These anti-τ_{had} candidates are τ_{had} candidates which have failed the identification criteria. The idea is that the kinematics of the jet is largely independent of the quality of the identification score. The anti-τ_{had} candidate is defined such that the candidate *fails* the identification criteria required for τ_{had} candidates but still passes a BDT score of 0.5. This minimum requirement rejects events with very low BDT score which are nearly entirely gluon initiated jets. These events are unrepresentative of the composition in the signal region.

In order to estimate the size of the background in the signal region, this template must have a event-by-event weight applied known as the "fake factor". This factor accounts for

the different probabilities of passing/failing the identification criteria ($FF \approx$ pass ID / fail ID).[4] The fake factor is dependent on a number of properties of the candidate including the number of charged tracks, the p_T of the τ_{had} candidate and the type of particle initiating the jet.

Separate fake factors are measured in bins of p_T for one pronged and three pronged τ_{had} candidates to account for the fake factors dependence on these properties. The type of particle, quark or gluon, which initiates the jet, affects the probability of the jet to be reconstructed as a (anti-)τ_{had}. Quark-initiated jets are often more likely to mimic τ_{had} candidates. Two processes contribute the most for these backgrounds: W+jets processes which produce an enriched sample of quark-initiated jets and multijet processes which are more gluon-enriched. A weighted fake factor is used which accounts for the expected relative contribution of W+jets and multijet of a template. The overall fake factor used to reweight events can be described by the equation:

$$FF = FF_{QCD}R_{QCD} + FF_{W}(1 - R_{QCD}), \tag{5.2}$$

where FF_i are fake factors for the indicated background process and R_{QCD} is the relative size of the expected multijet contribution in the template. Misidentified τ_{had} from other sources, such as top backgrounds, were estimated in MC to be negligible with respect to producing signal-like events from a misreconstructed τ_{had}.

The individual fake factors, comprising the overall FF, are measured in two separate background enriched control regions (CR). A W+jets enriched region is defined by requiring low transverse momentum and is used to measure the FF for a quark-enriched region. A multijet enriched region, derived by inverting lepton isolation requirements, measures the ratio of probabilities for gluon-enriched region. The fake factors range from 0.15-0.25 (0.01-0.04) for events with 1p (3p) τ_{had}.

The relative contribution is estimated by calculating the expected multijet contribution to a template with a data-driven technique. To estimate the number of multijet events an idea similar to the fake factor (recall equation 5.1) but with isolation requirements inverted rather than identification. A template of multijet events is derived from a region requiring an anti-τ_{had} *and* a lepton *failing* gradient isolation. Inverting the lepton isolation requirement enriches the region with jets producing a lepton from a secondary decay of a hadron: a very pure multijet region. This template is weighted in a similar manner to the fake factor but with a ratio known as the *isolation factor* ($IF \approx$ pass isolation / fail isolation).[5] The isolation factor is measured in a region where the final state particle candidates from the Higgs boson have the same charge. Overall the R_{QCD} tends to vary by up to 50% and is

[4]For brevity, subtraction of MC of events with a real τ_{had} from the numerator and denominator are omitted.

[5]For brevity, subtraction of MC of events with a real lepton from the numerator and denominator are omitted.

dependent on a number of kinematic factors such as the lepton p_T.

5.6.2 Top and Diboson Backgrounds

Several other backgrounds exist where the τ_{had} candidate is correctly identified but the kinematics of the event differ quite substantially. These include diboson, top and top-pair productions, all of which are modelled primarily with simulation. Of these, the largest background is the top and top-pair production. The diboson production, being rather small is simply taken from simulation.

Due to the large multiplicity of final states, the signal's signature can be easily replicated by top events. Given the large cross-section of the process, the overall size of the top background is non-negligible and thus a correct normalisation is required to adequately model this background. For such a process, a relatively pure control region can be defined and the normalisation of the background can be derived in a simultaneous fit of signal and control regions. Two control regions were defined in section 5.5.2 and the fit will be described later in section 5.9.

The remaining minor backgrounds contribution is from $Z \to ll$ + jets events could potentially mimic the signal process (particularly in the electron channel). This proceeds via a lepton misidentified as a τ_{had} rather than a jet and was found to be negligible so is modelled from MC simulation.

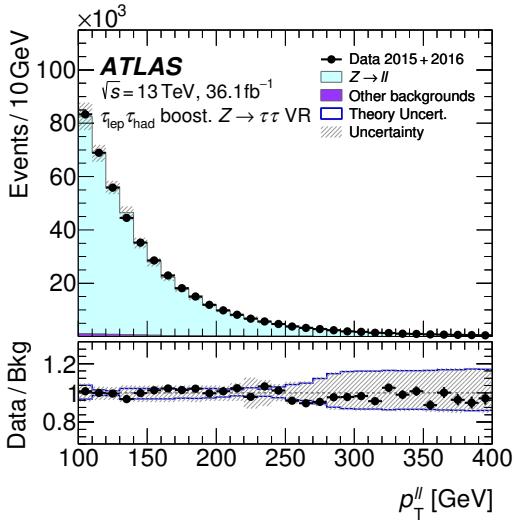
5.6.3 $Z \to \tau\tau$ Background

The only irreducible background is the $Z \to \tau\tau$ background which dominates the background composition in all signal regions. With an identical signal topology, differing mainly in the mass and spin of the Higgs/Z boson, this background is most difficult to remove while retaining a high signal acceptance. The modelling of the $Z \to \tau\tau$ also provides some of the most challenging systematic uncertainties to control.

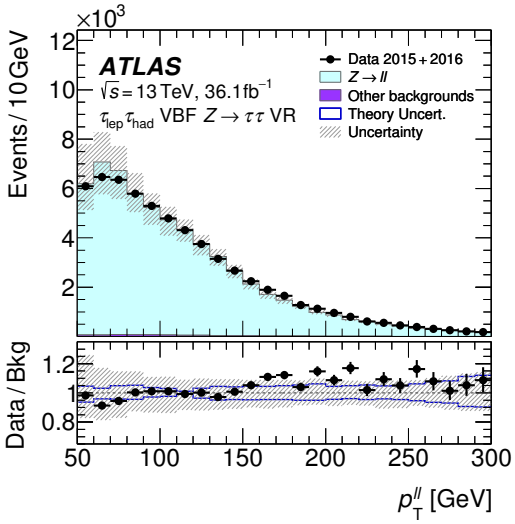
In the analysis of the $\sqrt{s} = 8$ TeV dataset, the embedding method [136] was employed. This method utilised observed $Z \to \mu\mu$ events to model the $Z \to \tau\tau$ background, replacing the muons with simulated τ_{had} decays. In using events from data, this ensured that the rate of $Z \to \tau\tau$ production as well as kinematic properties of the Z boson was modelled accurately. Being data-driven, there is no uncertainty in the Z boson kinematics. The only systematic uncertainty is due to the non-trivial procedure of replacing reconstructed muons with simulated τ_{had} calorimeter deposits [136]. Changes to the tracking reconstruction prior to data-taking at $\sqrt{s} = 13$ TeV now prevent the use of the embedding method. The approach used to model the $Z \to \tau\tau$ in this measurement is to take a MC estimate for the $Z \to \tau\tau$ shape but to normalise the background to data during the signal extraction procedure.

The MC estimate of the kinematic properties of the Z boson was validated with respect to events in two $Z \to ll$ + jets enriched regions (one for Boosted, one for VBF events). These $Z \to ll$ + jets enriched regions comprise dilepton events events where the dilepton mass greater than 80 GeV. Kinematic selections close to those defining the inclusive Boosted and VBF inclusive categories are applied to match the phase space of the signal regions. These regions are more than 99% pure in $Z \to ll$ + jets events. As the kinematics of the Z boson are unbiased with respect to the decay mode of the Z boson, $Z \to ll$ events can be used as a proxy for $Z \to \tau\tau$ events. Observables correlated with the mass estimate are shown in figure 5.7 to have good agreement between data and MC simulation. These observables include the lepton p_T, or variables relating to categorisation or to the kinematics of the Z boson.

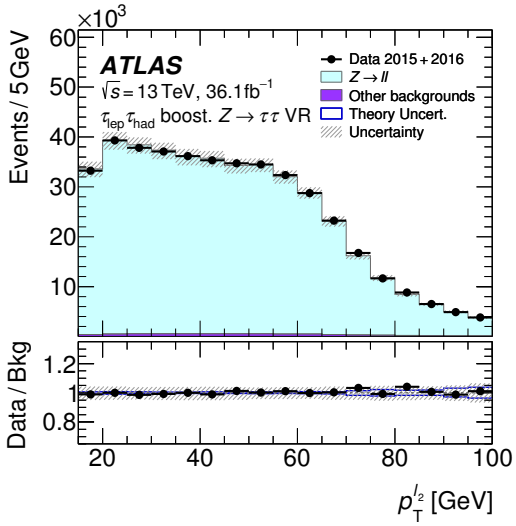
The size of the $Z \to \tau\tau$ background was derived from the signal extraction procedure, allowing the fit in the signal regions to constrain the normalisation of this background. The normalisation was calculated separately for VBF-like events and boosted ggH-like events as this impacts the Z production mechanism. Each normalisation is correlated amongst all three channels since the modelling of the Z boson should be independent of decay mode.



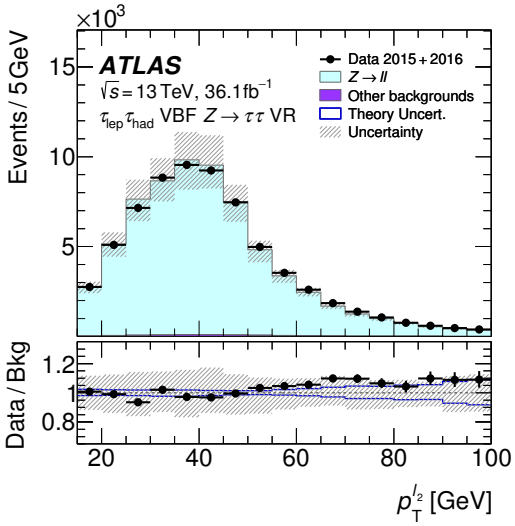
(a) p_T^Z (Boosted)



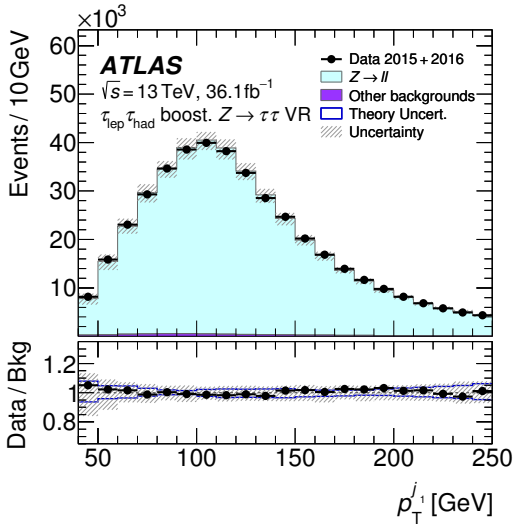
(b) p_T^Z (VBF)



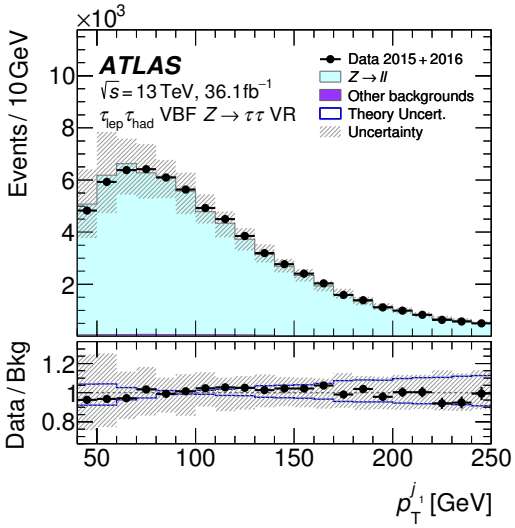
(c) $p_T(\ell)$ (Boosted)



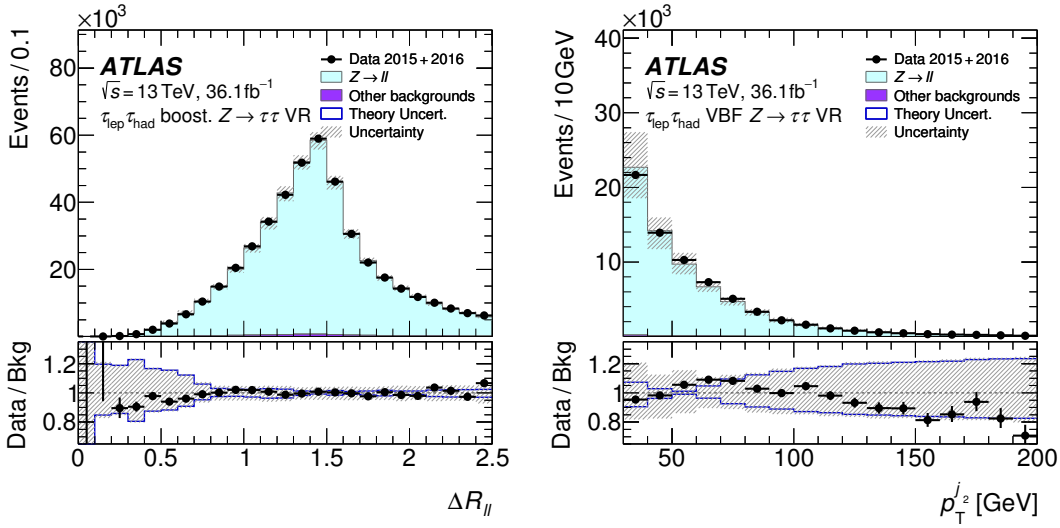
(d) $p_T(\ell)$ (VBF)



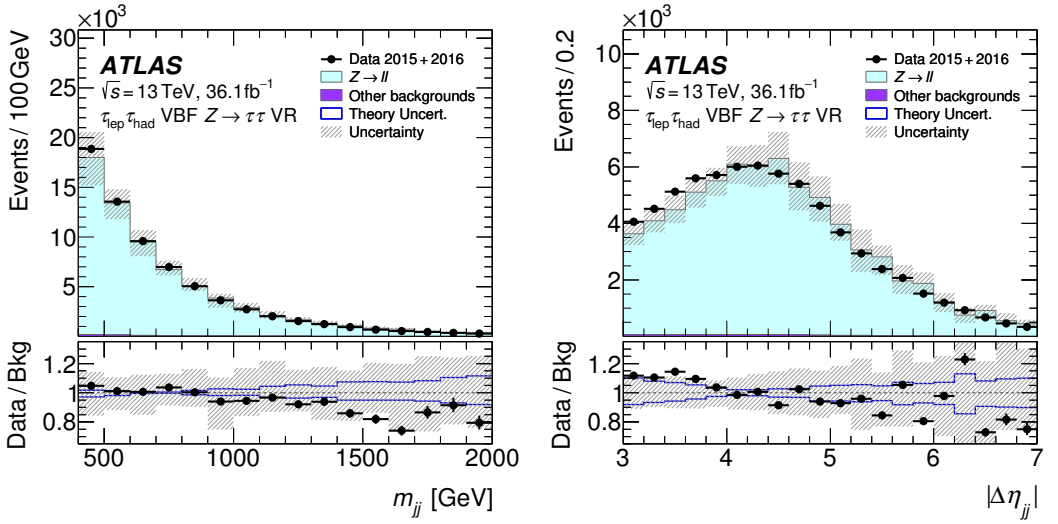
(e) $p_T(j_1)$ (Boost)



(f) $p_T(j_1)$ (VBF)



(g) $\Delta R(\ell_1, \ell_2)$ (Boosted) (h) $p_T(j_2)$ (VBF)



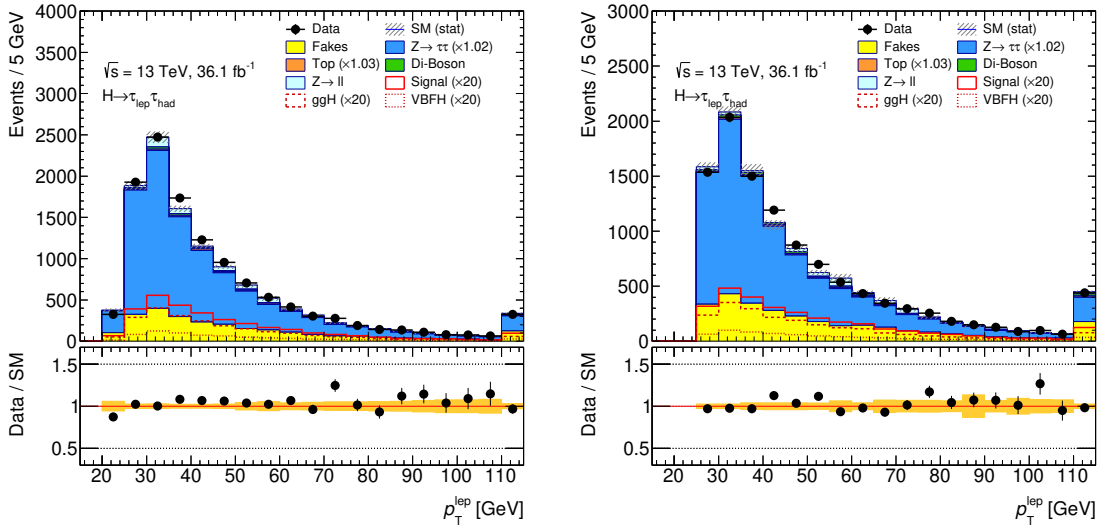
(i) m_{j_1,j_2} (VBF) (j) $\Delta\eta(j_1, j_2)$ (VBF)

Figure 5.7: Plots of various key distributions for the validation of the $Z \to \tau\tau$ in the $Z \to ll$ + jets CR corresponding to the semileptonic channel. Indicated in brackets is the specific $Z \to ll$ + jets CR used. Note p_T^Z is the p_T of the dilepton system. Here, the distributions are normalised to the event yield in data. Both statistical and systematic uncertainties. The theoretical uncertainty is highlighted independently in the dark blue line.

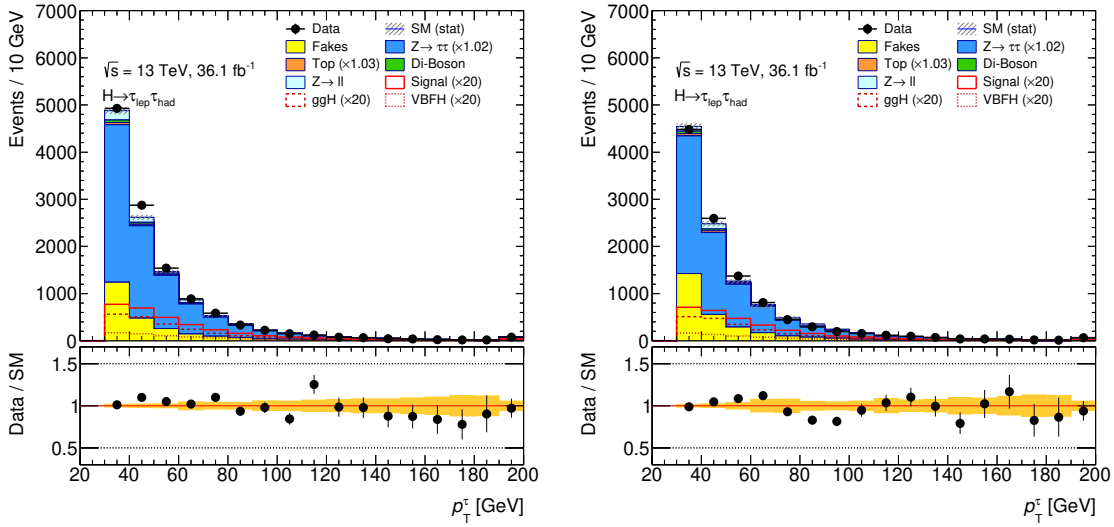
Theoretical systematic uncertainties will be addressed in section 5.8.2, however only the effect on the shape of the key discriminator (m_{MMC}) will be considered as the normalisation is decided by the fit. A scheme is implemented in order to properly account for events migrating between signal regions as a systematic uncertainty.

5.6.4 Modelling after Background Estimation

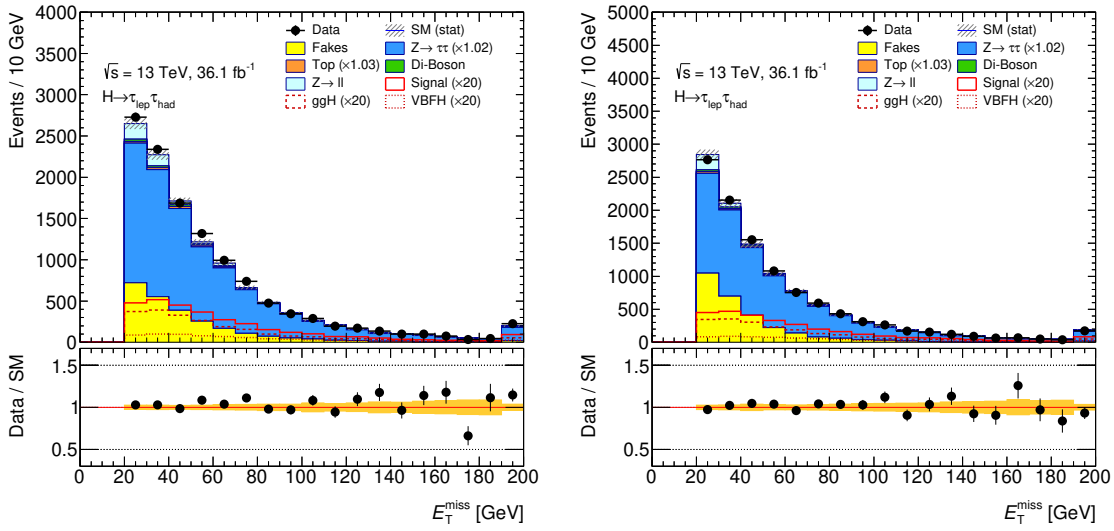
With the background estimations described here, variables were checked in several regions to assess the quality of the modelling. Distributions of some key variables demonstrating the modelling of the background estimate shown in figure 5.8 for the preselection. These build confidence that fundamental kinematics of the Higgs decay products as well as the number of jets, which is used to classify events, are modelled well by the background estimate. Key variables in separating the categories into signal regions are displayed in figures 5.9 and 5.10 for the boosted and VBF categories, respectively. The plots exemplify that the modelling of variables used to define the signal regions is fair.



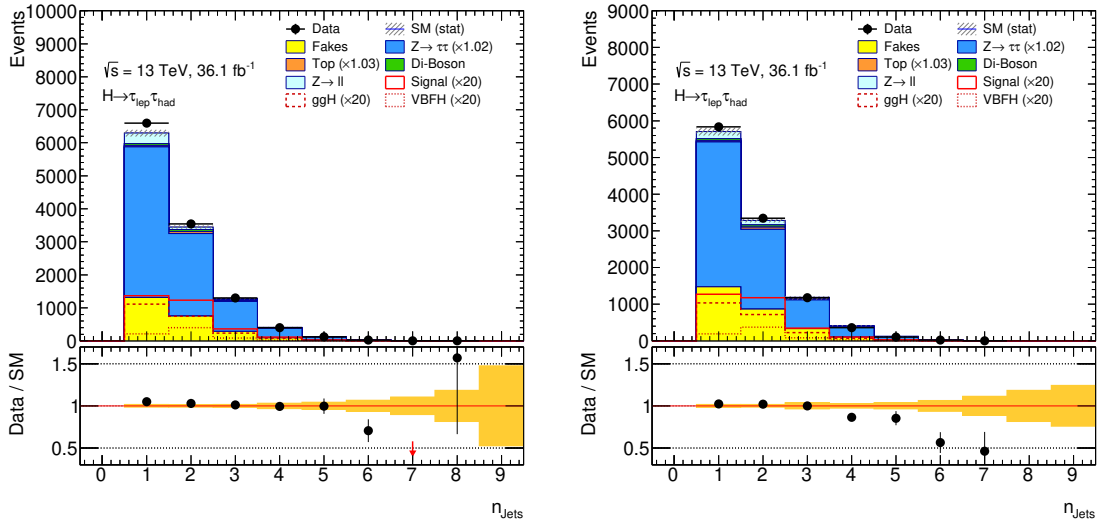
(a) $p_T(\ell)$



(b) $p_T(\tau_{had})$

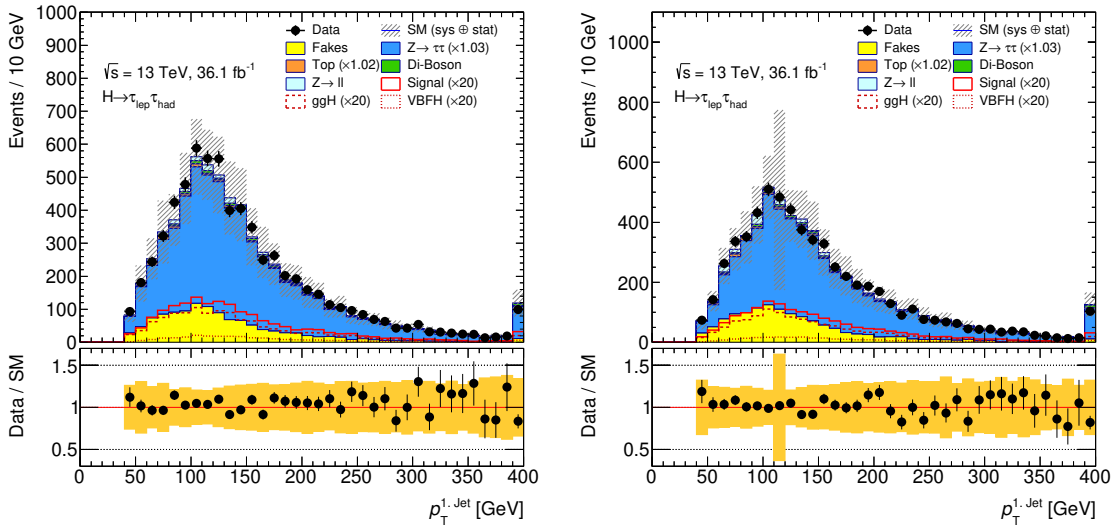


(c) E_T^{miss}

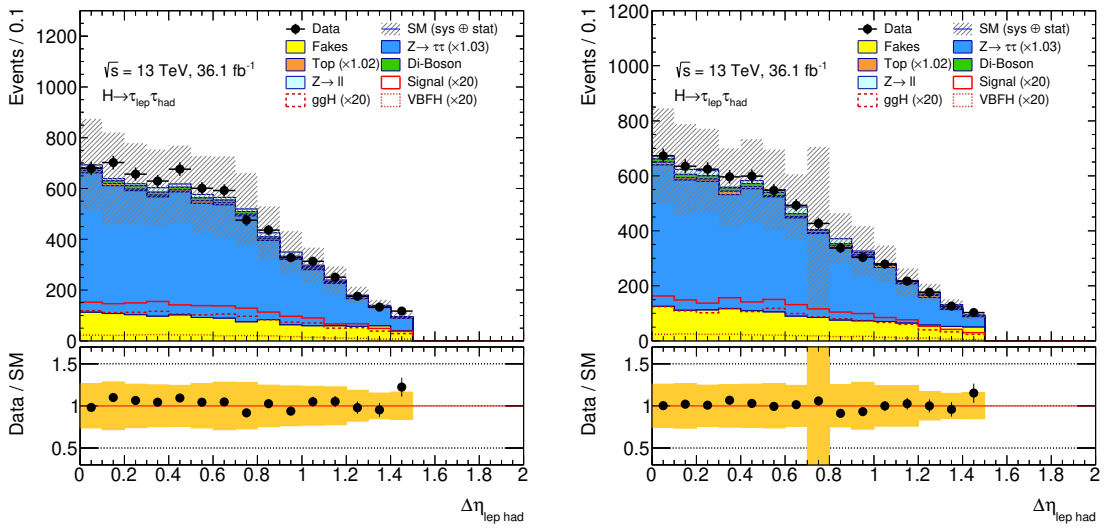


(d) Number of jets

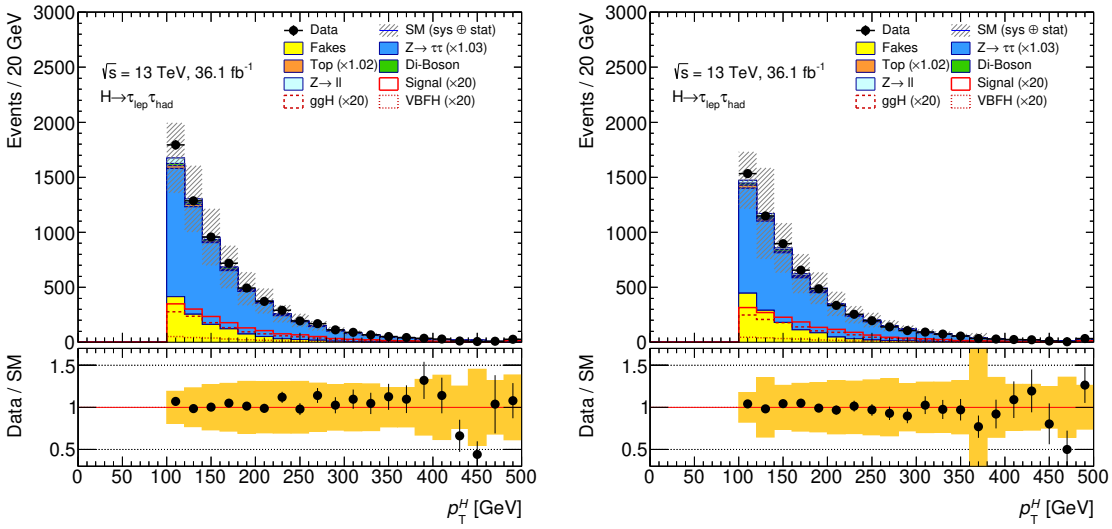
Figure 5.8: Plots of the preselection region in the muon and electron channels in the left and right columns respectively. The uncertainty band contains the quadratic sum of statistical and systematic uncertainties.



(a) Leading jet p_T

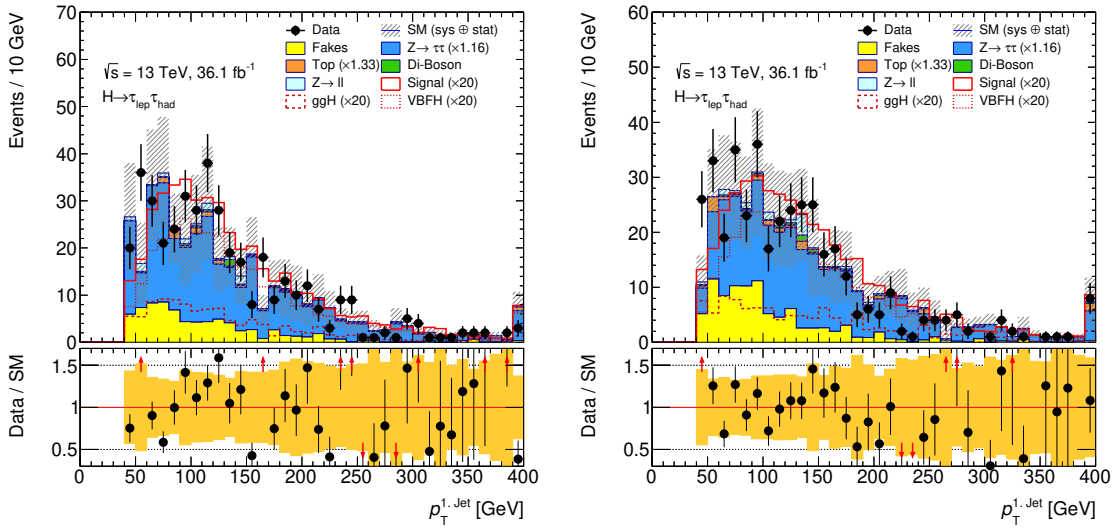


(b) $\Delta\eta(\ell, \tau_{had})$

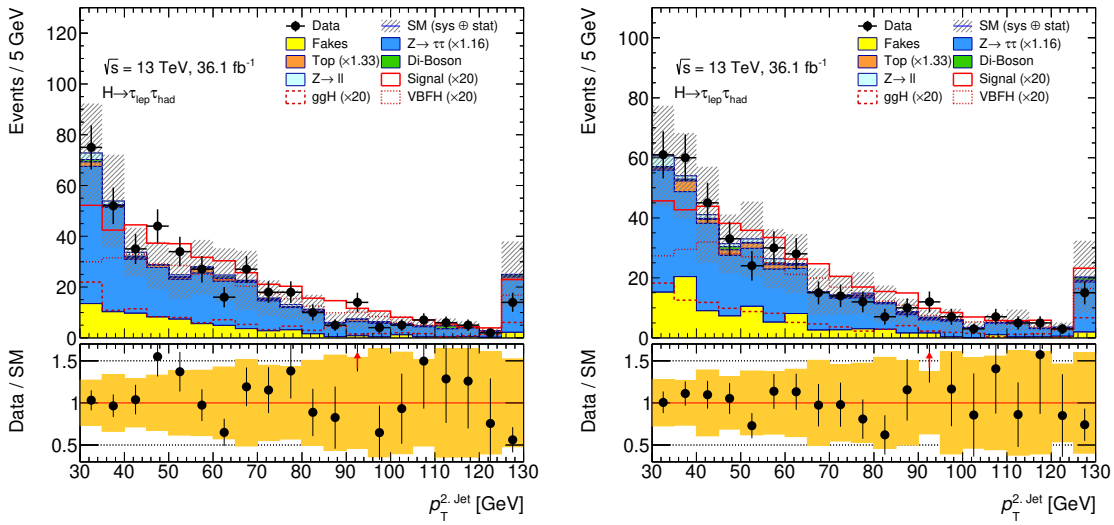


(c) p_T^H (Boost of the Higgs boson)

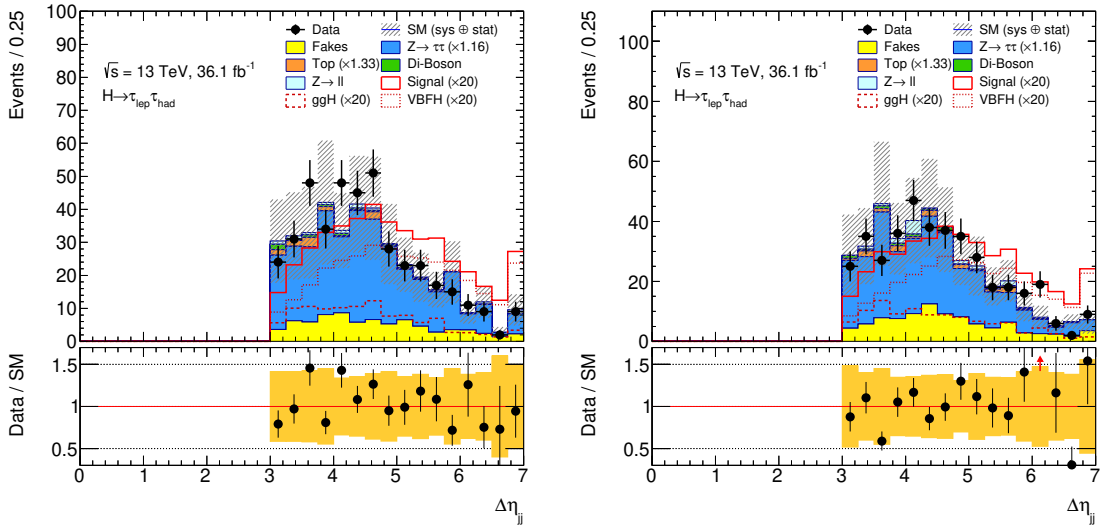
Figure 5.9: Plots of the inclusive boosted region in the muon and electron channels in the left and right columns respectively. Here, p_T^H represents the boost of the Higgs boson system in the transverse plane. This variable is calculated by summing the lepton, τ_{had} and E_T^{miss}, then taking the transverse component. The uncertainty band contains the quadratic sum of statistical and systematic uncertainties.



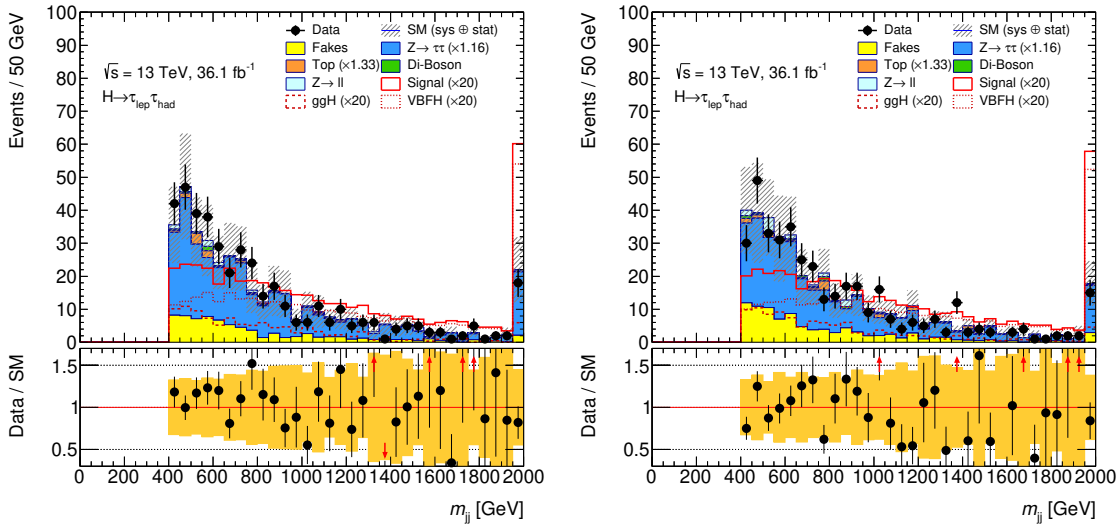
(a) Leading jet p_T



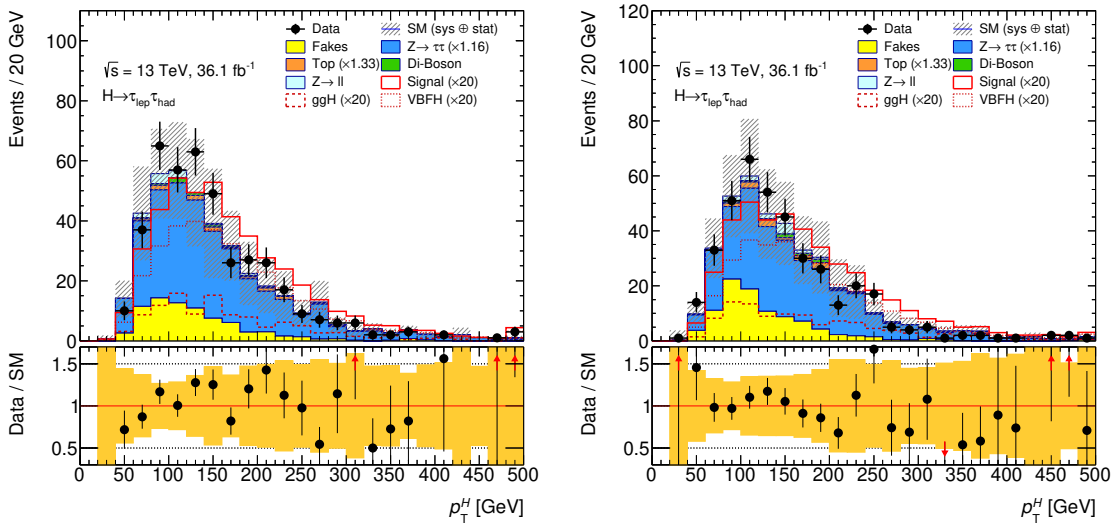
(b) Sub-leading jet p_T



(c) $\Delta\eta(j_1, j_2)$



(d) Dijet mass (m_{j_1,j_2}))



(e) p_T^H (Boost of the Higgs boson)

Figure 5.10: Plots of the inclusive VBF region in the muon and electron channels in the left and right columns respectively. Here, p_T^H represents the boost of the Higgs boson system in the transverse plane. This variable is calculated by summing the lepton, τ_{had} and E_T^{miss}, then taking the transverse component. The uncertainty band contains the quadratic sum of statistical and systematic uncertainties.

5.7 Higgs Boson Mass Reconstruction

From the plots in figures 5.9 and 5.10, it is clear that there is still a large background component in signal sensitive regions. The key discriminating variable which distinguishes the signal from background is the mass of the ditau system. The difference in mass between the Z boson and Higgs boson provides the best separation between the signal process and the $Z \to \tau\tau$ and $Z \to \ell\ell$ backgrounds. The misidentified τ_{had} background is expected to be rather flat across the mass spectrum as it is a non-resonant background. Thus by fitting the mass spectrum, the signal can be extracted with maximal sensitivity.

One of the challenges of ditau decays is the presence of between two and four neutrinos. The presence of a large number of neutrinos hinders the complete reconstruction of the Higgs boson mass. Given the four momenta of the neutrinos are not directly measurable, eight degrees of remain unconstrained for the given final state. After considering the τ lepton mass is fixed, the neutrino is approximately massless, and measuring the E_T^{miss}, only two degrees of freedom remain in order to properly reconstruct the Higgs boson mass.

Approximations are required to remove the remaining degrees of freedom in order to estimate the Higgs boson mass. Several approaches achieve this including the collinear mass approximation and missing mass calculator which will be discussed in the following sections.

5.7.1 Collinear Mass Approximation

If the missing energy is assumed to originate solely from the neutrinos in the τ decay, the kinematics of the $H \to \tau\tau$ system is described by the following set of equations:

$$
\begin{aligned}
E_x^{miss} &= p_{miss,1} \sin\theta_{miss,1} \cos\phi_{miss,1} + p_{miss,2} \sin\theta_{miss,2} \cos\phi_{miss,2} \\
E_y^{miss} &= p_{miss,1} \sin\theta_{miss,1} \sin\phi_{miss,1} + p_{miss,2} \sin\theta_{miss,2} \sin\phi_{miss,2} \\
m_\tau^2 &= m_{miss,1}^2 + m_{vis,1}^2 + 2\sqrt{p_{vis,1}^2 + m_{vis,1}^2}\sqrt{p_{miss,1}^2 + m_{miss,1}^2} - 2p_{vis,1}p_{miss,1}\cos\Omega_{vis,miss,1} \\
m_\tau^2 &= m_{miss,2}^2 + m_{vis,2}^2 + 2\sqrt{p_{vis,2}^2 + m_{vis,2}^2}\sqrt{p_{miss,2}^2 + m_{miss,2}^2} - 2p_{vis,2}p_{miss,2}\cos\Omega_{vis,miss,2}
\end{aligned}
$$

where $miss, i$ and vis, i indicate the missing and visible components of τ_i, and $\Omega_{vis,miss,i}$ is the angle between the visible and missing compontents of τ_i. This system has six unknowns when considering neutrinos as massless.

The collinear mass approximation is based on the assumption that the neutrino momenta are collinear with the visible products of the τ lepton. The E_T^{miss} vector can be expressed

as:

$$E_x^{miss} = p_{miss,1} \sin\theta_{vis,1} \cos\phi_{vis,1} + p_{miss,1} \sin\theta_{vis,1} \cos\phi_{vis,1}$$
$$E_y^{miss} = p_{miss,2} \sin\theta_{vis,2} \sin\phi_{vis,2} + p_{miss,2} \sin\theta_{vis,2} \sin\phi_{vis,2}$$

where $E_T^{miss}(x)$ and $E_T^{miss}(y)$ are the x and y components of the E_T^{miss} vector, and mis and vis refer to the missing (neutrino) and visible components of a τ decay, respectively. This approximation is made with the assumption $\phi^{mis} \simeq \phi^{vis}$ and $\theta^{mis} \simeq \theta^{vis}$.

The system is now fully solveable for the four-momenta of neutrinos and so the mass can be given by

$$m_{coll} = \frac{m_{vis}}{\sqrt{x_1 x_2}} \tag{5.3}$$

where x_i are the fraction of the τ momenta carried by the visible components of the τ

$$x_i = \frac{p_{vis,i}}{p_{vis,i} + p_{miss,i}}. \tag{5.4}$$

These x_i are the same outlined the event selections in section 5.5.

This approximation will not be used as a final discriminant in this chapter, however, will feature in the following chapter when considering lepton flavour violating Higgs boson decays in chapter 6.

5.7.2 Missing Mass Calculator

Another approach to constraining the missing degrees of freedom is to take a likelihood approach, known as the Missing Mass Calculator (MMC) method [137]. Instead of making one underlying assumption to estimate the kinematics of the neutrinos, the MMC method attempts to build likelihood functions in order to estimate the most likely value of the missing components.

These likelihood functions are constructed from distributions of the angles between visible τ decay products and the whole τ vector. The scan is then performed over the angles between visible and invisible components to maximise the likelihood function.

The MMC is the primary discriminating variable which is used in this measurement as it forms a distinctive peak for $H \to \tau\tau$ events around 125 GeV with a RMS of approximately 20 GeV. For this measurement, the MMC estimate is preferred over the collinear mass approximation mass (m_{coll}) as it shows less degradation in mass resolution in cases where the Higgs boson system is less boosted. The key discriminating variable for the measurement of the $H \to \tau\tau$ cross-section is the MMC mass (m_{MMC}) as it provides the best discrimination

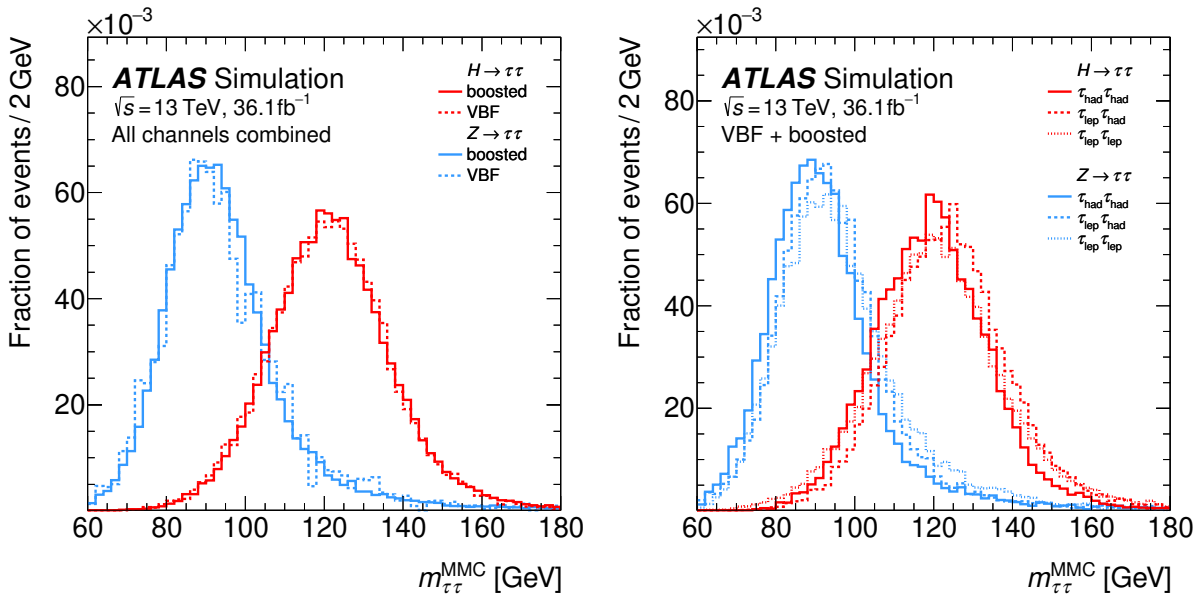


Figure 5.11: Distribution of the Higgs boson mass, reconstructed with the MMC algorithm split by category (left) and channel (right)

against the $Z \to \tau\tau$ background while the multijet background, being non-resonant, is flat against the m_{MMC} variable.

The distribution of the MMC masses are presented per channel and category are presented in figure 5.11. Here it is evident a clear separation between $Z \to \tau\tau$ and $H \to \tau\tau$ events is possible with this reconstructed mass variable. With a well defined discriminator, the next step is to evaluate systematic uncertainties and perform hypothesis testing.

5.8 Systematic Uncertainties

Before performing the statistical tests required to measure the coupling, systematic uncertainties affecting the measurement must be addressed. Numerous sources of systematic uncertainties exist which can be categorised as experimental uncertainties (involving the calibration of particle candidates), theoretical uncertainties (relating to the choice of theoretical models used to generate the MC) and background modelling uncertainties (due to the procedure chosen to model data-driven backgrounds). Uncertainties in the data-driven backgrounds are described in section 5.6.1.

5.8.1 Experimental Uncertainties

Particle candidate reconstruction involves several stages of calibration which result in a number of sources of systematic uncertainties. Uncertainties relating the reconstruction are discussed in chapters 3 and 4 however a short summary is presented below.

Efficiency Uncertainties

In reconstructing and identifying particle candidates, the efficiencies of reconstruction, identification and triggering are measured and corrected using measurements such as the tag-and-probe methods previously discussed. Uncertainties associated with the measurement/correction of efficiencies are taken from these measurements and propagated through the analysis to determine the effect of these types of systematic uncertainties.

Uncertainties for identification and reconstruction for electrons [138], muons [85] are rather small contributing, on average, $< 3\%$ to the expected yield. For τ_{had} [97] the impact due to identification efficiency corrections range from 2-6%.

The uncertainties in the flavour tagging (in)efficiency [92] produce effects which are typically $< 1\%$. For jet reconstruction, the uncertainty for measuring the reconstruction efficiency and JVT tagging efficiency are also $< 1\%$ in their effects on yields.

Energy Scale and Resolution Uncertainties

The measurement and correction of the energy/momentum scales for reconstructed particles provides another source of uncertainty. The energy scale measurements can impact not only the normalisation (as the efficiency does) but also the shape of key variables.

The variations due to the insitu calibration TES [97] are previously discussed and contribute 1% or less. Other TES variations impact at most 4% in certain regions. The energy scale variations for jets [89] due to many factors, including flavour composition, pileup effects and η dependence, are included additionally. These have a broad range of impact on the normalisation from between 1 and 10%. Further variations for electron [139] and muon

energy scales [85] are also assessed although all effects are rather low (less than 1%).

Uncertainties relating to the resolution of energy scales are also included for jets and electrons as well as the reconstructed E_T^{miss}. Uncertainties due to electron energy resolution [139] is negligible but E_T^{miss} resolution uncertainties [140] contribute a 1% variation on the overall yield. The jet energy resolution (JER) [89] is of particular importance for this measurement. The poor resolution of jets leads to very large smearing of the E_T^{miss}, ultimately propagating to the important m_{MMC} discriminator. As a result the JER uncertainty is parameterised into 11 uncorrelated sources rather than combining into one global uncertainty. The effect on the normalisation of background samples reaches as high as 24%. This is most pronounced in the VBF phase space and causes degradation in the resolution of the m_{MMC} distribution.

Luminosity and Pileup

The uncertainty on the measured luminosity, as measured by dedicated luminosity detectors adjacent to the ATLAS detector [80], is a flat 2.9% across all MC samples. Uncertainty due to the pileup reweighting process contributes between 1-3%.

5.8.2 Theoretical Uncertainties

As processes at the LHC predominantly are produced from QCD interactions, the modelling of the production is dependent on theoretical calculations and experimental inputs. Limitations due to calculation of QCD at fixed order/scale require assessment for systematic effects.

Variations due to different parton showering models are also accounted for by comparisons to alternative generators. The procedure follows recommendations from the LHC Higgs Cross Section Working Group [21].

Signal uncertainties

Uncertainties affecting the total cross-section of the generated Higgs signal are the QCD scale, PDF and α_S values used. One nuisance parameter is assigned for the total cross-section uncertainty and other variations are added as uncorrelated parameters affecting shape and acceptance.

The uncertainty due to the truncation of the QCD expansion for VBF and VH production modes is assessed by varying the renormalisation and factorisation scales between 0.5 to 2.0 times the value used in simulation. The largest envelope formed from these variations is taken as the uncertainty for the QCD scale for these processes. The ggH production mode requires a more detailed treatment due to selections made with regard to the number of jets introducing large corrections.

Variations from the factorisation and renormalisation scale are taken but are supplemented with variations in:

- The resummation scale and migrations between the 0, 1 and 2 jet categories [141],
- Variation in the VBF phase space as well as effects of a third jet veto [26],
- The dependence of the Higgs p_T shape,
- The dependence on the top mass.

The variations are treated as uncorrelated and are factorised into components affecting: the total cross-section, the impact on the shape of the discriminant and the relative impact on the acceptance of events into individual SR. The factorisation for the sources of uncertainty due to theoretical calculations follow the same division.

Variations arising from the use of a specific PDF are determined according to recommendations from the PDF4LHC [142] and number 30 in total. The contribution of the uncertainties on the experimentally determined α_S are also accounted for.

Finally the effect of utilising the Pythia8 parton shower is checked by comparing to the Herwig7 parton shower and underlying event uncertainties are evaluated according to recommendations from the LHC Higgs Cross Section Working Group [21].

Overall, the uncertainty due to the QCD scale is the dominant theoretical effect, reaching as high as 10% on the normalisation of the signal cross-section, with significant contributions from PDF uncertainties to a similar degree.

$Z \to \tau\tau$ uncertainties

The uncertainties for the $Z \to \tau\tau$ background is driven by the following uncertainties:

- The renormalisation and factorisation scales,
- The PDF used,
- The CKKW [143] scheme which matches jet to parton,
- The resummation scale (denoted as QSF),
- Underlying event uncertainty,
- The parton shower modelling.

The renormalisation and factorisation scales as well as the PDF uncertainties are given by the Sherpa generator. The CKKW and QSF uncertainties are calculated with truth information as a function of the p_T of the Z boson as well as the number of jets. Underlying event

uncertainties were found to be negligible so were not included in the fit model. The parton shower uncertainty is calculated by comparing the SHERPA and MADGRAPH generators [114].

The total normalisation of the $Z \to \tau\tau$ is left as a floating parameter determined by the fit so the total uncertainty on the $Z \to \tau\tau$ is not a nuisance parameter as in the case of the signal samples. Instead, the effects of these variations are taken as two sets of components. One set is defined solely as shape variations impact the shape but also account for migrationary effects between SR within a category (e.g. events migrating from VBF loose to VBF tight due to a given variation) and is given per variation, category and channel. Another set is defined to account for differences between the inclusive category definitions between channels (e.g. the difference in normalisation due to the PDF in the Boosted category) which is taken as a normalisation uncertainty and correlated across channels.

The largest component of the theoretical uncertainties affecting the $Z \to \tau\tau$ normalisation is the CKK matching of jet to parton. This contributes between 2-5% depending on the region. QCD scale uncertainties are a subleading effect contributing up to 4% to the $Z \to \tau\tau$ normalisation.

Background Modelling Uncertainties

The modelling of misidentified τ_{had} background process is affected by a number of systematic uncertainties. In the FF calculation, the four components to the final FF value are the FF_{QCD}, FF_{W}, R_{QCD} and the anti-τ_{had} events.

The uncertainty due to statistical uncertainty on the individual fake factors, FF_{QCD} and FF_{W}, are treated as a single source of uncertainty for either boosted or VBF topologies. The uncertainty is most pronounced in the VBF region where the effect is of order 30-40%, with the effect in the boosted region at approximately 9%.

The calculation of R_{QCD} is also limited by considerations of limited statistics in both data and MC. Systematic effects which impact the R_{QCD} are rather small ranging from 1-5%. The statistical uncertainty in the anti-τ_{had} regions are also taken as a systematic uncertainty on the total number of predicted fakes.

Finally a systematic uncertainty is assigned to account for differences observed in the closure test in which the procedure is repeated but inverting the roles of the OS and SS requirements. The difference between prediction and observed data in the SS regions are taken as a uncertainty on the template for the OS region. This effect ranges from 2-5% on the normalisation of the jet fake background.

5.9 Statistical Hypothesis Testing

With a well modelled set of signal regions and rigorously measured systematic effects, the next step is to derive a result. This involves a statistical hypothesis testing involving building a model, and testing with respect to observations.

A test statistic encoding the predictive model is initially defined. This model incorporates all relevant channels, regions and correlations between signal/background processes and systematic corrections/uncertainties. The model is fit with respect to the data in order to estimate many of the normalisations of the $Z \to \tau\tau$ and top backgrounds which are not set pre-fit. The observation will then be interpreted one of two ways: either as a statistical significance of observing an excess, or an exclusion of certain signals.

The first few parts of this section will describe how a test statistic is defined. Following this is a description of how an observation is interpreted within the statistical framework used by the ATLAS collaboration. Finally the specific statistical model and fit procedure used in this measurement will be presented.

5.9.1 Test Statistic

A measurement is performed by measuring an observable such as a mass variable or BDT discriminant. The distribution of the discriminant is binned and for each bin (indexed by i) the expected number of events is described by:

$$E[n_i] = \mu s_i(\theta) + b_i(\theta) \tag{5.5}$$

where s_i and b_i are the expected signal and background content in bin i respectively. Either μ represents the *parameter of interest* (POI) for the measurement. In this measurement the POI is either the signal strength, which represents the signal with respect to the SM expectation (σ_{obs}/σ_{SM}) or it is the cross-section of the $H \to \tau\tau$ process $\sigma_{H\to\tau\tau}$ (the formula is modified slightly in this case). The expected number of events is often affected by statistical and systematic uncertainties which are parameterised in θ.

The events in each bin are assumed to follow a Poissonian distribution (ignoring effects from systematic uncertainties). Including systematic effects involves modifying the likelihood with additional terms. Thus the likelihood function can be defined as:

$$L(\mu, \theta) = \prod_{i \in \text{bins}} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \times L_{aux}(\theta). \tag{5.6}$$

where n_i is the total observed value, θ is a term representing the systematic uncertainties and L_{aux} encodes this into a likelihood. The systematic uncertainties are parameterised multiplying by extra terms (also bin by bin) which are called "nuisance parameters". The

effect of nuisance parameters (NP) are profiled in a fit in order to minimise the likelihood function. Each NP is treated in a specific manner befitting the type of systematic producing the effect. The NPs can have both a shape and normalisation component to them. The shape and normalisation components are dealt with separately as they have specific issues with both.

For most NPs the shape components add a Gaussian multiplicative factor whereas the normalisation component (including theoretical uncertainties) utilises a log-normal distribution, which prevents potential negative content bins from causing issues in the fit procedure. A Gaussian term is additionally used for the luminosity uncertainty. For statistical uncertainty due to the background model, a gamma probability distribution is utilised.

According to the Neyman-Pearson lemma, the optimal test statistic can be constructed by the profile likelihood ratio

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(\hat{\mu}, \hat{\theta})} \tag{5.7}$$

where the numerator is the conditional maximum likelihood and the denominator the unconditional maximum likelihood. The conditional maximum value is determined from with a conditional fit, where θ maximises the likelihood for a fixed value of μ. The unconditional maximum value is determined from a fit where both μ and θ are allowed to vary independently in order to maximise the likelihood.

The value of μ are not inherently constrained to be positive so the test statistic needs to be defined as

$$q_\mu = \begin{cases} -2\log\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(\hat{\mu},\hat{\theta})}, & \hat{\mu} \geq 0 \\ 0, & \hat{\mu} < 0, \end{cases} \tag{5.8}$$

in order to account for this as negative values are possible. This is the test statistic used for hypothesis testing.

5.9.2 Hypothesis Testing

With a test statistic defined, one can test a hypothesis. One of two goals are sought. If an excess above the null hypothesis exists, calculate the significance to claim a discovery. If no such excess exists, set an exclusion limit constraining a particular parameter.

Significance

The significance measures how unlikely it is to make a specific observation, given a null hypothesis. Starting with p-value, the probability of measuring a result as extreme as the

one observed (given a particular signal hypothesis μ), which is defined by

$$p_\mu = \int_{q_{0,obs}}^{\infty} f(q_\mu|\mu) dq_\mu \tag{5.9}$$

can be transformed to a significance by simply taking Gaussian intervals. The function f is the probability density function (q_μ) which is parameterised as a product of Poisson distributions (for the statistical uncertainty) and Gaussian distributions (for systematic uncertainties) over the number of bins of the distribution. According to the Ward approximation, p_μ is a monotonic function.

For background-only hypothesis corresponds to the case of $\mu = 0$ where the significance is calculated from p_0. By convention a discovery is claimed at 5σ or more (corresponding to a p_0 of 2.89×10^{-7}).

Exclusion Limits

In cases where a search is not expected to yield a discovery, the sensitivity of the search to a parameter is the figure of merit. In order to assess this, the limit at which a search can reveal a signal (termed exclusion limit) is determined. To start, the test statistic is modified to

$$q_\mu = \begin{cases} -2\log \frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(0,\hat{\theta}(0))}, & \hat{\mu} < 0 \\ -2\log \frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(\hat{\mu},\hat{\theta})}, & 0 \leq \hat{\mu} \leq \mu \\ 0, & \hat{\mu} \geq 0, \end{cases} \cdot \tag{5.10}$$

Following the CL_s prescription for method signals satisfying $\mathrm{CL}_s < 0.05$ where CL_s is defined as

$$\mathrm{CL}_s = \frac{p_\mu}{1 - p_b} \tag{5.11}$$

where p_b is the case where $\mu = 0$ (the background-only hypothesis). This approach is taken to alleviate issues regarding spurious signal and look else-where effects.

5.9.3 Statistical Model

The test statistic for this measurement is built upon predictions in all three channels (fully leptonic, semileptonic and fully hadronic) for both categories (Boosted and VBF). Also included are relevant CR for the normalisation of important MC based background estimates: in the case of the semileptonic channel this consists of two top enriched CR of a single bin.

The background samples included in each region are grouped as follows: $Z \to \tau\tau$, $Z \to \ell\ell$, misidentified τ_{had}, top backgrounds as well as "other" backgrounds too small to be considered as an independent sample. The signal samples were separated by production mode to

a ggH, VBF and VH+ttH sample, respectively.

Sources of uncertainty are parameterised as NPs and are correlated across samples and regions per source of systematic uncertainty. The exception to this rule is for the data-driven estimates of misidentified τ backgrounds, these are treated as uncorrelated per channel and the normalisation is fixed within this model.

As discussed previously (in section 5.6), the normalisation of the top and $Z \to \tau\tau$ backgrounds are left to be constrained by the fit. In both cases a single normalisation is correlated between the relevant sample in each of the boosted or VBF category separately. In the case of the top background in the semileptonic channel, this is correlated between the SR (in either boosted or VBF category) and a top enriched CR (in the corresponding category). For the $Z \to \tau\tau$ normalisation, the normalisation of the $Z \to \tau\tau$ (in either boosted or VBF category) is not correlated to any CR but is rather between the $Z \to \tau\tau$ sample in each of the SR across all three channels (in the corresponding category).

The entire systematic model is described by the schematic in figure 5.12.

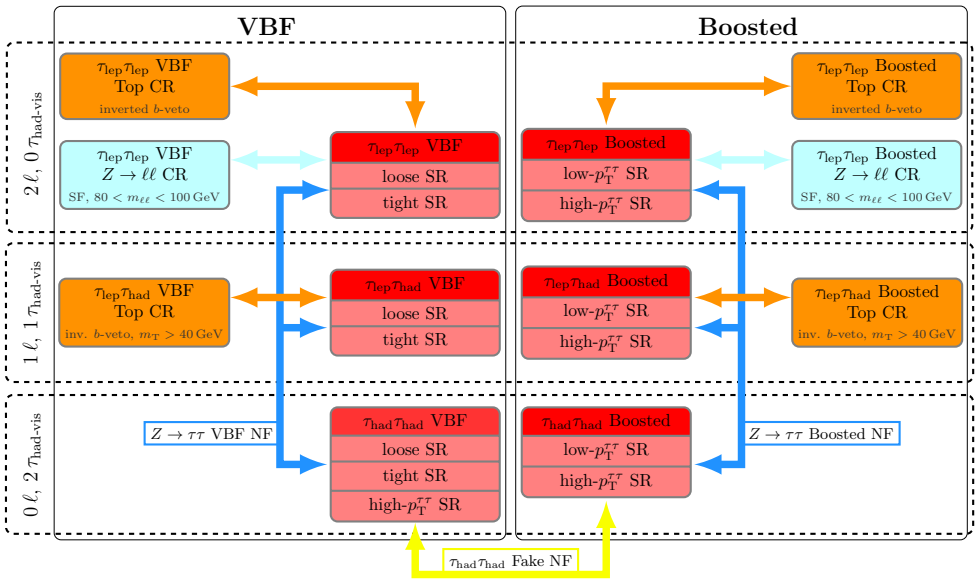


Figure 5.12: A schematic of the fit model. The signal regions (red) are divided between the VBF (left) and Boosted (right) categories and the three channels (fully leptonic, semileptonic and fully hadronic from top to bottom). Orange and blue boxes represents the top and $Z \to \ell\ell$ enriched control regions used within the fit. Coloured arrows represents the correlations of the normalisation of certain backgrounds which are floated within the fit.

5.9.4 Fit procedure

The fit performed is a binned profile likelihood fit. "Binned" refers to the overall likelihood function being a product of terms based on the bin contents m_{MMC}. "Profile" indicates the fit scans the NPs in order to reach a minima during the fit procedure. The nuisance parameters can be constrained or pulled by varying degrees during the fit. Constraints indicate the favourability of data to be able to assess the size of a variation. Pulls indicate that the data is distributed closer to the $\pm 1\sigma$ variation rather than the nominal.

There are more than 200 independent NPs, including all experimental, theoretical systematics as well as normalisations which are floated for top and $Z \rightarrow \tau\tau$ backgrounds, The complexity of such a fit, coupled with NPs with a high degree of statistical fluctuation or insignificant NPs, can cause instabilities in the fit which lead to non-converging fits. To simplify the model, certain criteria are used to reduce the number of nuisance parameters (known as pruning) which ultimately do not impact the likelihood model due to the small size of the variation or the noisiness of the shape. The four criteria used to prune the model are as follows:

- **Overal normalisation**: Nuisance parameters affecting the normalisation are only accepted into the fit if the total integral (normalisation) of the variation differs from the nominal's by at least 0.5%.
- χ^2 **test**: This criteria affects the shape component of the NP. A requirement that at least one of the upward/downward variation fails the criteria that their reduce χ^2 is larger than 0.1.
- **Statistical uncertainty**: Also affecting the shape component, large statistical uncertainties in the variation can create shapes which are essentially statistical noise. If the entire distribution has no significant variation, then no information is added to the fit. Hence if the overall statistical uncertainty is less than 0.1 of the integral of the variation, the shape component is removed.
- **Significance testing**: Another means by assessing the importance of the shape component is to consider whether the variation has a significant effect in at least one bin. A measure is defined as $S_i = |u_i - d_i|/\sigma_i^{nom}$ where u_i and d_i are the variation in bin i and σ_i^{nom} is the statistical uncertainty on the nominal value. If no bin has $S_i > 0.1$, the shape component is removed.

Effects from such variations may also hamper the fit's convergence. Thus the shape components of the variations are also smoothed in order to ensure little noise prevents the model from fitting well. This smoothing is performed on the ratio between the variation and the nominal as this minimises artefacts which can be created by the smoothing.

Occasionally one sided variations on a bin-by-bin basis may still be present which leads to further convergence issues. In cases such as this, a further symmetrisation criteria is

applied on a bin-by-bin basis. If in a bin, a variation is one-sided with respect to the background it affects, the largest of the up and down variation (with respect to the nominal) is mirrored about the nominal in order to produce a symmetric variation (in that particular bin).

In the case where the variation is negative (either before or after symmetrisation) the value of the variation is set to be 1×10^{-6} in order to have a well defined variation with respect to the fit model.

Initially the fit is tested against an Asimov dataset built from the sum of the expected signal and backgrounds in place of the observed data. This dataset is used to assess the stability of the fit model as well as provide an expected sensitivity given the fit model for analysis. When the data has power to constrain overestimated systematic effects, the NLL distribution from a fit to an Asimov dataset will show a reduced 1σ width with respect to the minima. If an NP is over-constrained by data (or constrained where there is no reason) the respective 1σ band will be less than what is input, indicating a need to scrutinise the fit.

A subsequent fit is then performed in which the data in signal sensitive bins remain fixed to the expected signal plus background hypothesis in the signal sensitive bins, but the sideband regions (those outside these bins) utilise real events for the data model. This fit is called a "hybrid fit" as the data model is a hybrid between the expected data and real data. Pseudodata is created where events outside a mass window between 100 and 150 GeV are replaced with Asimov data. Fitting this pseudodata tests whether the data has a preference for NPs to be pulled away from their nominal values, whilst still being unbiased with respect to signal-sensitive data. In these cases, they can often indicate problems in the fit model and prompt studies to explain the reasoning behind this.

Both sets of fits were found to be sensible with no major pulls or constraints in any nuisance parameters. Following the validation of the fit to Asimov and hybrid data, the fit to the real data is performed and the results derived from such a fit.

5.10 Results

This section will detail the results of the fit to data, first in the semileptonic channel and then the combination with other channels.

5.10.1 Semileptonic Channel

The post-fit distributions of the MMC mass in each of the signal regions in the semileptonic channel are presented in figure 5.13 and yields are summarised in 5.5.

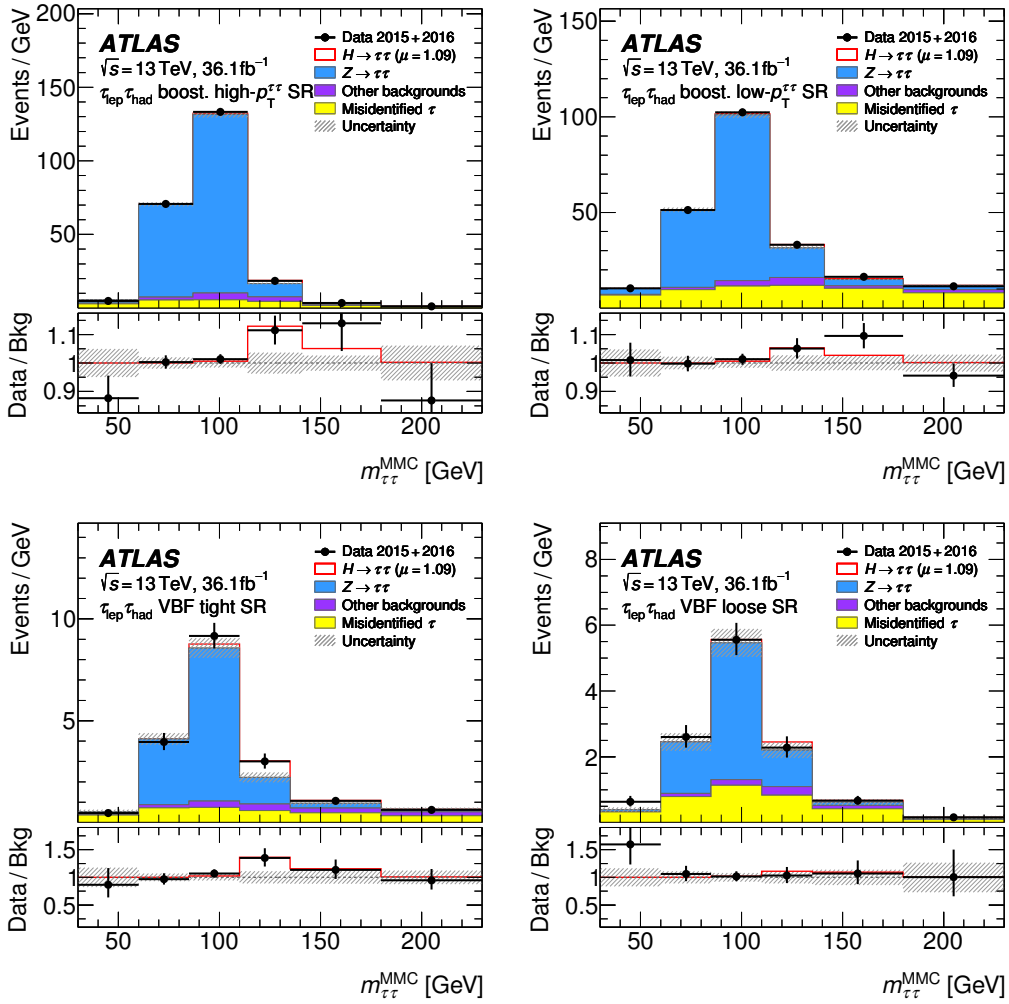


Figure 5.13: Post-fit distributions of the m_{MMC} discriminant for the semileptonic channel. Plots above show Boosted regions and below show the VBF regions with the higher $p_T^{\tau\tau}$ on the left. The lower panel shows the ratio of data over the background estimate. The uncertainty band represents the post-fit uncertainty which includes both statistical and systematic uncertainties.

	VBF				Boosted			
	Loose		Tight		Low-p_T		High-p_T	
$Z \to \tau\tau$	175	18	319	22	4159	96	5313	92
$Z \to ll$ + jets	10.1	3.0	12.6	3.0	130	37	115	16
Top	5.8	1.6	17.9	4.7	119	20	56	10
Misidentified τ_{had}	103	16	100	15	1907	77	617	27
Other backgrounds	4.0	1.6	9.5	1.9	115.2	7.9	129.6	8.9
ggH, $H \to \tau\tau$	4.1	1.3	7.0	2.0	62	16	64	21
VBF, $H \to \tau\tau$	7.5	2.2	25.3	7.1	12.0	3.5	14.2	4.1
WH, $H \to \tau\tau$	< 0.1		0.1	0.0	4.0	1.1	5.3	1.4
ZH, $H \to \tau\tau$	< 0.1		< 0.1		1.8	0.5	2.8	0.8
Total background	299	18	459	23	6430	88	6230	92
Total signal	11.7	3.3	32.5	8.4	80	20	86	24
Data	318		496		6556		6347	

Table 5.5: The yields calculated by the fit in all the semileptonic SR. Note the second column for each SR is the overall uncertainty including systematic uncertainty.

An excess in the semileptonic channel, above background-only null hypothesis is observed with a significance of 2.4σ (2.6σ expected). A breakdown of the significance of the excess observed in each of the boosted and VBF regions is presented in table 5.6. These reveal that the Boosted channel is driving the sensitivity due to the larger uncertainties in the VBF categories. This was found to be consistent with other channels.

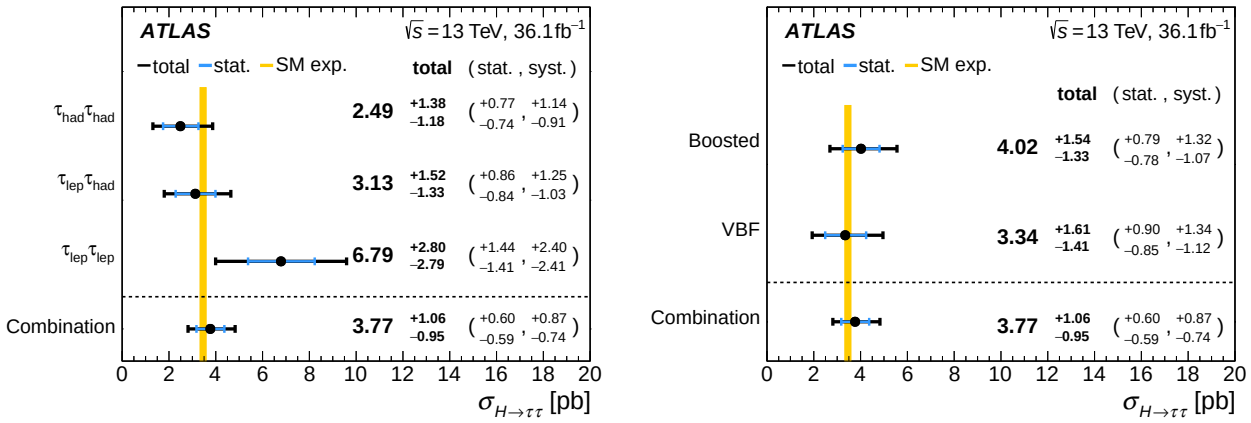
	$\tau_{lep}\tau_{had}$	Boosted	VBF
Median signficance	2.6	1.8	1.70
Median p0	0.0043	0.033	0.044
Observed signficance	2.4	1.9	1.65
Observed p0	0.0077	0.030	0.050

Table 5.6: The expected and observed significances of signal for the combined channel as well as the boosted and VBF categories.

The best-fit signal strength for the semileptonic channel is $0.9^{+0.44}_{-0.38}$, compatible with the SM expectation (of 1). Fits to the boosted and VBF categories separately in the semileptonic channel both demonstrated consistency with the SM expectation.

5.10.2 Combination

The combination of all decay channels with 36.1 fb^{-1} reveals an excess above the no Higgs null hypothesis of 4.4σ (with an expected value of 4.1σ). The best-fit signal strength for the combination of all channels is observed to be $1.08^{+0.31}_{-0.28}$, very much compatible with the SM. The best-fit cross-section $\sigma_{H\to\tau\tau}$ for each the decay channels as well as the combined boosted and combined VBF categories are presented in figure 5.14. In contrast to the measurement with the $\sqrt{s} = 8$ TeV dataset, the boosted category now leads the sensitivity over the VBF region.



(a) Breakdown per channel

(b) Breakdown per category

Figure 5.14: Best fit values for the cross-section for the combined channel fit divided into either channels or categories. The yellow vertical band represents the SM expectation of 3.46 ± 0.13 pb. Blue bands indicate the statistical only uncertainty whereas black indicates the total.

The semileptonic channel remains as one of the strongest channels in this measurement, on par with the fully hadronic channel with regard to the uncertainty on $\sigma_{H\to\tau\tau}$. The fit was found to be stable and well behaved, no extreme pulls or constraints were observed in the leading sources of uncertainty. The overall ranking of nuisance parameters based on their impact on the final fitted cross-section for the combined fit is summarised in 5.15.

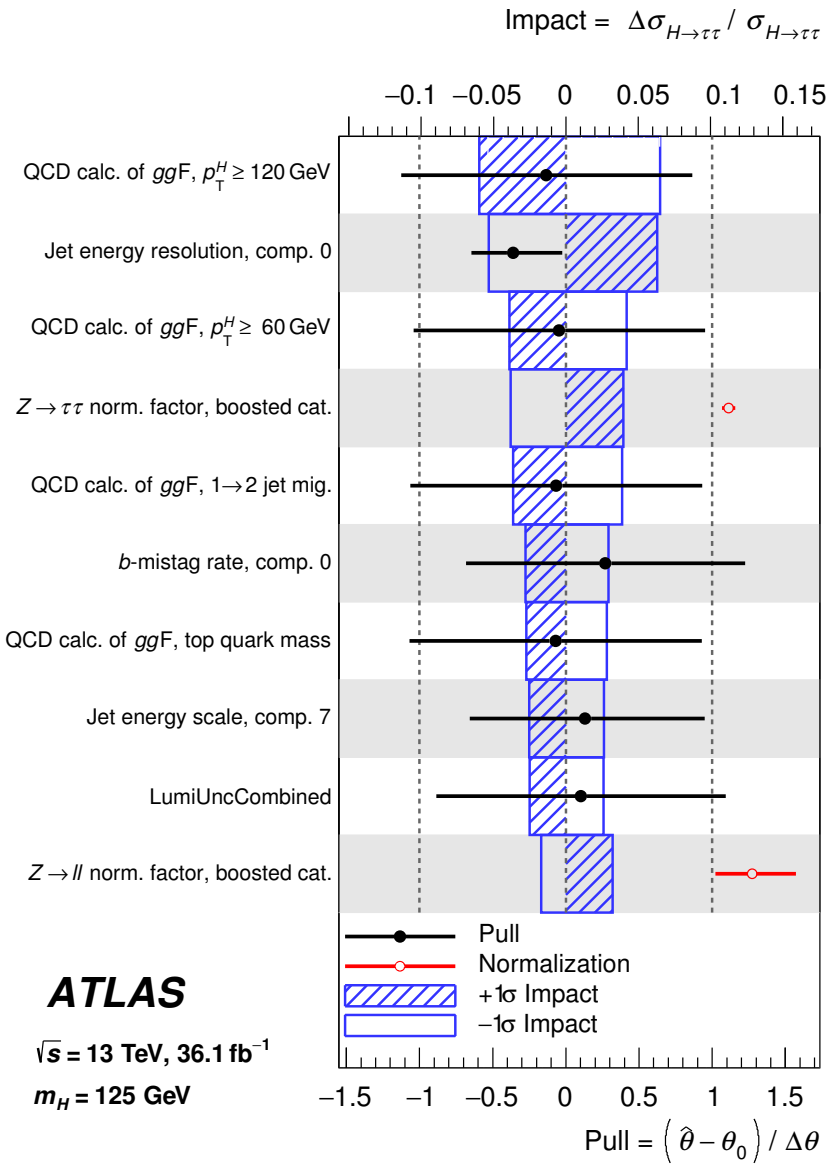


Figure 5.15: The leading 25 nuisance parameters ranked by the impact on the best-fit cross-section (blue). The black points show the pull of the nuisance parameter.

The top ranked NPs are dominated by jet, E_T^{miss}, theory and background estimation systematics. The jet and E_T^{miss} uncertainties impact the reconstruction of neutrino momenta in calculating the key m_{MMC} discriminator. Large corrections for QCD also produce substantial effects on the ggH produced signal. Further background estimation systematics, in particular from the normalisation of the $Z \to \tau\tau$ background also play a large role in this measurement. A full breakdown of sources of systematics is presented in table 5.7. The normalisation of

Source of uncertainty	Impact $\Delta\sigma/\sigma_{H\to\tau\tau}$ [%]			
	Observed		Expected	
Theoretical uncert. in signal	+13.5	-8.7	+11.9	-7.7
Background statistics	+11	-10	+10.2	-9.8
Jets and E_T^{miss}	+11.5	-9.3	+10.5	-8.6
Background normalization	+6.8	-4.8	+6.6	-4.6
Misidentified τ_{had}	+4.5	-4.2	+3.7	-3.4
Theoretical uncert. in background	+4.6	-3.6	+5.1	-4.2
Hadronic τ_{had} decays	+4.7	-3.0	+5.8	-4.2
Flavor tagging	+3.3	-2.4	+2.9	-2.2
Luminosity	+3.3	-2.3	+3.1	-2.2
Electrons and muons	+1.2	-1.0	+1.1	-0.9
Total systematic uncert.	+24	-20	+22	-19
Data statistics	±16		±15	
Total	+28	-26	+27	-25

Table 5.7: Expected and observed impact on the $\sigma_{H\to\tau\tau}$ by various sources of uncertainty. The impact is expressed as a fraction of the best fit $\sigma_{H\to\tau\tau}$.

the $Z \to \tau\tau$ was determined to be $1.04^{+0.10}_{-0.09}$ and 1.11 ± 0.05 times the predicted value for the VBF and boosted categories respectively. The normalisation of the $Z \to \tau\tau$ in the boosted category produced a particularly large impact on the uncertainty for the total cross-section. For the top background the fit determined the normalisation to be $1.53^{+0.30}_{-0.27}$ and 1.12 ± 0.07 times the predicted value for the VBF and boosted categories respectively.

The cross-section from this measurement is $\sigma_{H\to\tau\tau} = 3.71 \pm 0.59(\text{stat})^{+0.87}_{0.74}(\text{syst})$ pb which can be compared to a predicted SM cross section of $\sigma^{SM}_{H\to\tau\tau} = 3.43 \pm 0.13$ pb.

An additional two-parameter fit on the cross sections of $\sigma^{ggH}_{H\to\tau\tau}$ and $\sigma^{VBF}_{H\to\tau\tau}$ determined to be $3.0 \pm 1.0(\text{stat})^{+1.6}_{-1.2}(\text{syst})$ and $0.28 \pm 0.09(\text{stat})^{+0.11}_{-0.09}(\text{syst})$ pb, respectively. The 2D contour plot of the best fit $\sigma^{ggH}_{H\to\tau\tau}$ and $\sigma^{VBF}_{H\to\tau\tau}$ values is presented in figure 5.16. This is consistent with SM expectations of $\sigma^{ggH}_{H\to\tau\tau} = 3.0 \pm 0.13$ and $\sigma^{VBF}_{H\to\tau\tau} = 0.237 \pm 0.006$ pb, respectively.

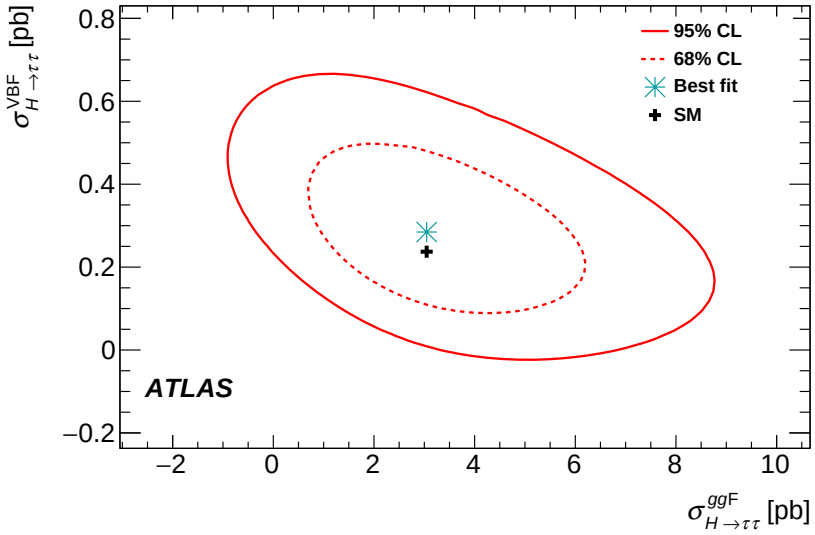


Figure 5.16: Best fit values for the cross-sections of the ggH and VBF production modes.

In combination with analyses performed on $\sqrt{s} = 7$ and 8 TeV, the total excess observed with data collected to the end of 2016 yields a significance of 6.4σ observed (5.4σ expected). This constitutes a discovery of the $H \to \tau\tau$ decay process with the ATLAS detector, complementing the aforementioned CMS result.

5.11 Conclusion

The first measurement of the $H \to \tau\tau$ decay process was made with ATLAS combining the available datasets from ATLAS measurements up until the end of 2016. The measurement is consistent with the SM expectation with a cross-section measured as $\sigma_{H\to\tau\tau} = 3.70 \pm 0.58(\text{stat})^{+0.89}_{-0.75}(\text{syst})$ pb. The semileptonic channel has been shown to be one of the two leading channels driving the sensitivity of this measurement, especially after the reoptimisation of the event selections. Results from this measurement will subsequently be combined with CMS measurements to provide global measurements of cross-sections.

In addition to the total cross-section, a simplified template cross-section is expected to be measured, providing theorists with a global picture of the status of Higgs measurements in interesting regions of phase space. In future, this will be followed by a full differential cross-section measurement, providing the most accurate breakdown of the $H \to \tau\tau$ measurement.

6

Search for Lepton Flavour Violating Higgs Boson Decays

The work from chapter 5 has established the Higgs boson's coupling to third generation leptons. Along with measurements of the coupling to the third generation of quarks, this has firmly established that the 125 GeV resonance is indeed the Higgs boson predicted from the SM. With the exception of couplings to light fermions and the Higgs boson's self couplings, all of which require a substantially larger dataset to measure, all SM couplings of the Higgs boson have been measured. The focus of Higgs boson physics now shifts to searching for new phenomena in Higgs boson decays.

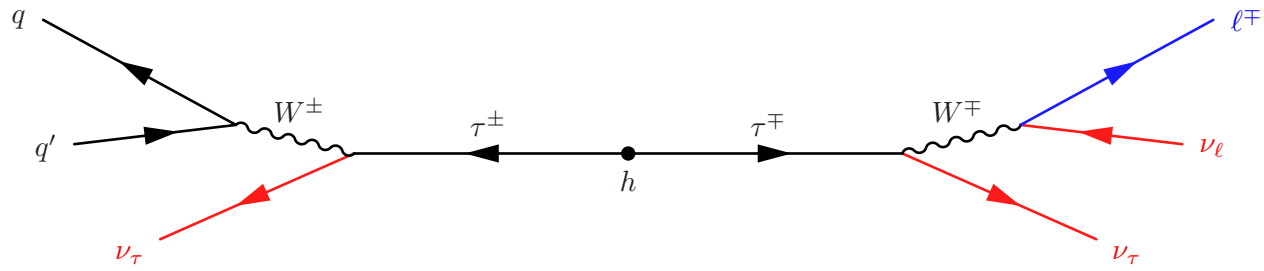
One such phenomena of interest is charged lepton flavour violation. The search for LFV in Higgs boson decays is motivated by theoretical considerations presented in section 2.4.2 including excesses in $\sqrt{s} = 8$ TeV datasets [68, 71, 72]. Charged lepton flavour violation within the Higgs sector would indicate non-zero off-diagonal Yukawa couplings between the Higgs boson and leptons. The search for LFV Higgs boson decays with $\sqrt{s} = 13$ TeV datasets is expected to improve the sensitivity as there is an enhancement of the Higgs boson production cross-section and a larger dataset (36.1 fb^{-1}) allowing for either confirmation of the $\sqrt{s} = 8$ TeV result or stronger limits to be placed.

Constraining off-diagonal elements is possible by searching for decays of the Higgs boson to a pair of leptons with different flavours; $H \to e\mu$, $H \to e\tau$ and $H \to \mu\tau$. As the upper limit on $\mu \to e\gamma$ decays are very strong [63], a direct $H \to e\mu$ search is not expected to provide better sensitivity to the $Y_{e\mu}$ coupling. Thus only the $H \to e\tau$ and $H \to \mu\tau$ searches will be pursued. Amongst the $H \to l\tau$ decays, the case where the τ lepton decays to τ_{had} is more sensitive than the case of fully leptonic decays. This is due to the larger branching fraction for τ lepton decays to hadrons.

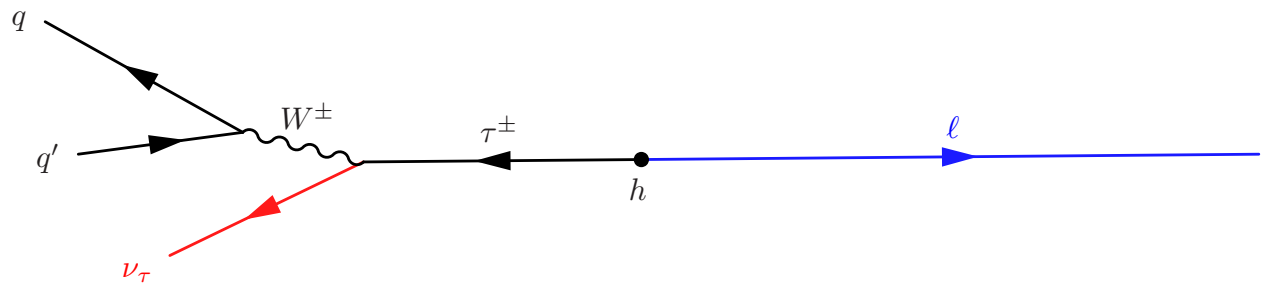
A search for both $H \to e\tau$ and one for $H \to \mu\tau$ is presented in this chapter. These searches are separate as constraints from $\mu \to e\gamma$ allow the presence of only one of the $H \to \ell\tau$ decays. Several methods from chapter 5 are invoked to perform this search.

6.1 Signal and Background

This chapter focuses on the $H \to e\tau$ and $H \to \mu\tau$ decays in the case where the τ decays to hadrons due to the high branching fraction to hadrons. The signal topology closely mimics that of the semileptonic $H \to \tau\tau$ decays. A comparison of the $H \to e\tau$ and $H \to \mu\tau$ processes and the semileptonic decays of $H \to \tau\tau$ is sketched in figure 6.1.



(a) Semileptonic SM di-τ decay



(b) LFV decay to $\ell^{\mp}\tau^{\pm}_{had}$

Figure 6.1: Topology of a SM Higgs boson decay compared with a LFV Higgs boson decay. This is not a Feynman diagram but a schematic in the Higgs boson rest frame to indicate the kinematics of the decay. Indicated in blue is the source of the lepton and in red are the sources of neutrinos.

The visible signature (a light lepton and τ_{had} candidate) is identical in both cases, they differ only in the number of neutrinos present in the final state. In the case of the LFV decays, the final state contains two less neutrinos as the lepton does not originate from a τ lepton decay.

The backgrounds for this search are inclusive of those described in section 5.1.2 with the addition that the $H \to \tau\tau$ process becomes a source of background. Major backgrounds include the $Z \to \tau\tau$ and misidentified τ_{had} backgrounds with smaller contributions from top, diboson and $H \to \tau\tau$ production.

The $Z \to \tau\tau$ background can be distinguished from the LFV signal by considerations of

the missing energy. The missing energy is directed along the τ_{had} four-vector in the case of the LFV decay whereas in ditau decays the direction is not so focussed. In addition, the presence of two fewer neutrinos gives the visible mass good discrimination. Misidentified τ_{had} backgrounds are also a substantial background due to the large cross-section of production.

Data-driven backgrounds will described in 6.5, primarily for the misidentified τ_{had} background. Remaining backgrounds are based on MC simulation which was presented in the previous chapter in table 5.2. The signal was simulated with the same generators as the $H \to \tau\tau$ background but with the PYTHIA generator simulating the decay to one lepton and one τ_{had}.

6.2 Analysis Strategy

Two approaches are devised to extract the signal: a cut-based (CBA) and a multivariate (MVA) approach. The cut-based approach applies sequential selections on sensitive variables to build a number of signal regions. Conversely, the multivariate approach attempts to utilise a larger phase space, with a high signal acceptance, but gains sensitivity from a dedicated classifier. The multivariate approach achieves its signal sensitivity by training a classifier which takes into account the correlations of many variables. Taken individually, these variables may not provide little discrimination between signal and background but with a trained MVA is able to outperform a CBA. The MVA is used as the primary approach for the derivation of the final results. The CBA method is based on a similar approach to that of the analysis of $\sqrt{s} = 8$ TeV data [71, 72] and was only considered as a cross-check. The result with CBA was consistent with the results MVA but with a weaker upper limit.

Each search is divided into two "channels", which depends on the decay mode of the τ lepton. The hadronic decay mode, having a much larger branching fraction, is the more sensitive channel and was the focus of the author's work. Each channel is divided into a VBF and non-VBF category. The VBF category targets the VBF signal whereas the non-VBF category contains events failing this criteria and thus contains most of the signal. This chapter will focus on the non-VBF category.

This search utilises 36.1 fb^{-1} of $\sqrt{s} = 13$ TeV collision data. Single lepton triggers define the dataset used, requiring a lepton p_T threshold of 27.3 GeV and a gradient isolation. More details can be found in section 5.3.

6.3 Event Selections

An initial preselection is defined followed by two broad categories; a VBF category (targeting the VBF topology) and a non-VBF region (containing mostly the ggH-produced signal). As the key issue with detecting an LFV Higgs boson decay is due to the small production cross-section (already $\mathrm{BR}(H \to l\tau)) > 1\%$ have been excluded), the non-VBF category is expected to drive the sensitivity. Thus, the focus of this chapter is the non-VBF category (in the hadronic decay channel). The relevant selections are summarised in table 6.1.

The preselection selects one lepton and one τ_{had} candidate with opposite charge. The lepton is required to pass a medium identification and have a $p_T > 27.3$ GeV and with a gradient isolation to satisfy trigger requirements. The τ_{had} candidate must pass a *tight* working point. This works to reject the background from jets misidentified as τ_{had}, which contributes a large fraction to the overall background. With this working point, these backgrounds contribute still 45% in some signal regions. The τ_{had} must also have a $p_T > 20$ GeV and a electron BDT veto score corresponding to 95% signal acceptance for 1p τ_{had}.

A kinematic selection on $\Sigma\cos(\Delta\phi)$ rejects W+jets backgrounds whereas a selection on $|\Delta\eta(\ell, \tau_{had})|$ reduces multijet events. A b-jet veto is also required to reject top quark backgrounds.

Signal Regions

Among events passing preselection, the VBF category is defined by passing selections on the jet kinematics (in a similar manner to the last chapter). The non-VBF region is defined as events which pass the preselection but fail the VBF categorisation. The yields and background compositions are depicted in figure 6.2. The regions are composed of approximately half $Z \to \tau\tau$ events and between 33% to 40% of events with misidentified as a τ_{had}. The non-VBF region has high signal acceptance but also very poor signal purity, less than 1%. The signal sensitivity comes from the approach of signal extraction. Here a multivariate approach will be used to maximise the signal sensitivity. As the current constraints from CMS [68] are below 1% a low signal yield is expected. Thus it is prudent to use a approach which minimises signal loss but yet can still have good discrimination with respect to backgrounds.

	$e\tau_{had}$	$\mu\tau_{had}$
	Preselection	
Electron Veto BDT	Exactly one electron candidate No muon candidates $> 0.15(0.05)$ for 1p (3p) τ_{had}	Exactly one muon candidate No electron candidates -
Lepton charge	At least one τ_{had} candidate passing *tight* identification Lepton and τ_{had} are oppositely charged	
$\Sigma\cos(\Delta\phi)$	> -0.35	
$\lvert\Delta\eta(\ell,\tau_{had})\rvert$	< 2	
b-jets	No b-jet candidates	
	VBF	
τ_{had} p_T	Pass preselection > 45 GeV	
Number of jets	> 1 (with $p_T > 30$ GeV)	
Leading jet p_T	> 45 GeV	
m_{jj}	> 400 GeV	
$\lvert\Delta\eta_{jj}\rvert$	> 3	
	non-VBF	
	Pass preselection Fail VBF selections	

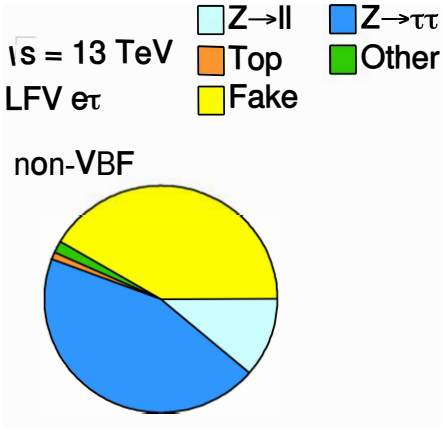
Table 6.1: Summary of selections for preselection and VBF categories for both the $e\tau_{had}$ and $\mu\tau_{had}$ channels. Note minimum selections on particle candidates were detailed in the previous chapter 5.4 unless specifically mentioned in section 6.3. The non-VBF category is defined as the events passing preselection but failing VBF categorisation.

Control Regions

The fake factor method involves measuring rates of events passing identification in background enriched control regions. For the non-VBF category, a W+jets enriched region is defined by inverting the requirement on $\Sigma\cos(\Delta\phi)$ and a multijet enriched region is defined by inverting the selection on $|\Delta\eta|$ and requiring $m_T(l, E_T^{miss}) < 60$ GeV.

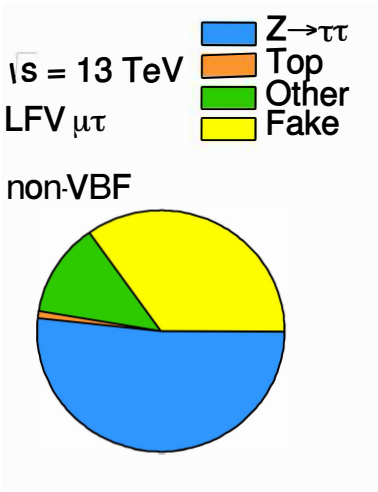
No dedicated top control region is required specifically for the τ_{had} channel due to it's negligible contribution. Dedicated $Z \to \tau\tau$ control regions, one for non-VBF and one for VBF, which are $\approx 80\%$ pure in $Z \to \tau\tau$ are used to normalise the $Z \to \tau\tau$ background. These control regions are defined with a dilepton selection from the leptonic decay channel where lepton p_T requirements are inverted.

Sample	Yield
$Z \to \tau\tau$	76360 ± 290
Fake	69600 ± 500
Top	1770 ± 26
$Z \to \ell\ell$	19400 ± 150
Others	27100 ± 1000
Background	171700 ± 700
Signal	1176.5 ± 3.5
Data	168883



(a) $e\tau_{had}$ non-VBF channel

Sample	Yield
$Z \to \tau\tau$	95170 ± 310
Fake	63090 ± 280
Top	1763 ± 26
Others	25500 ± 600
Background	184300 ± 700
Signal	1199.7 ± 3.5
Data	184508



(b) $\mu\tau_{had}$ non-VBF channel

Figure 6.2: Yields and background composition of the non-VBF categories. Note in the $\mu\tau_{had}$ search, the $Z \to ll$ + jets events are summed into the "Others" background whereas in the $e\tau_{had}$ search this is separated due to the dedicated background estimate for misidentified electrons. Uncertainties stated are only statistical. Signal yield is assuming BR of 1%.

6.4 Multivariate Approach

A multivariate approach was implemented using boosted decision trees (BDT) in order to better classify LFV signal from backgrounds.

6.4.1 Boosted Decision Trees

The BDT is a supervised learning algorithm aimed at classifying background and signal events utilising multiple input variables. The BDT conceptually begins as a simple decision tree in which events can be classified by performing sequential selections on a set of input variables. This is exemplified by the schematic in figure 6.3.

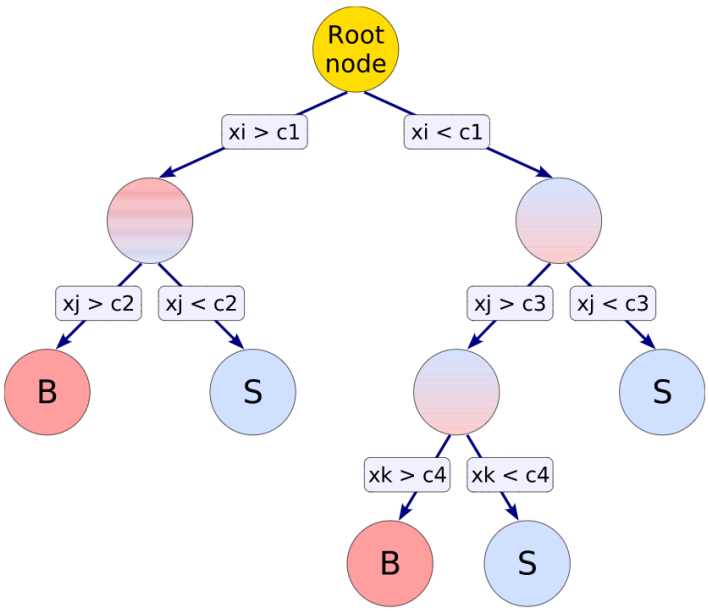


Figure 6.3: A example of a single decision tree.

The choice of selection is made by scanning the phase space and performing a selection which best separates signal and background. This process continues until a terminating condition is met. This decision tree would then have terminal nodes (or leafs) which can be labelled as "background" or "signal" nodes depending the signal to background ratio in the final node. The problem with a simple decision tree is that often it is not much more performant than the traditional CBA. This is because it replicates the structure of a CBA of sequential selections which maximise the signal sensitivity in a node/region. Thus, instead of relying on a single tree, multiple trees can be created as an ensemble method.

The method of *boosting* was introduced to increase the effectiveness of a single weak classifier by creating an ensemble which on the whole performs better. The *gradient* method of boosting used relies on creating new trees which minimise a given loss function (a measure of how correctly a tree classifies events). Iteratively, the residuals of the predictions from prior trees are used as inputs for the creation of the subsequent trees. Calculating the gradient from a differentiable function gives the path to the minima of the loss function, thus allowing for new subsequent trees to be more optimal. The collection of trees allows for a discrimination power which exceeds that of a single decision tree. A single combined classifier score is constructed from the classification power of the individual trees. Each terminal node per tree in the ensemble is given a weight and an ensemble score is calculated. This score represents a probability an event should be categorised as signal or background. The BDT score classifies events using this single ensemble score. Events more likely to be background being attributed scores closer to 0 and those likely to be signal closer to 1. For this analysis, a gradient boosting method is used in the package XGBoost [144].

One issue with boosting is overfitting: where an algorithm overperforms on a training dataset compared to the application to an independent dataset. This indicates the algorithm can not reliably be generalised to use on data. A process of "regularisation" is implemented in this algorithm to prevent overfitting by penalising more complex structures.[1]

The standard metric for the classification power of a multivariate algorithm is the integral of a receiver operating curve (ROC). The ROC curve measures the signal efficiency against the background efficiency of a classifier when selecting events above a certain threshold. The area under this curve (AUC) indicates the separation power of this classifier.

6.4.2 Inputs and Modelling Checks

The dataset used to train the relevant BDTs is derived from a looser version of the non-VBF region in which the τ_{had} candidate is required to pass the *medium* identification criteria (both signal and background samples). This loosening allows for a larger dataset to be trained upon which reduces problems due to overfitting. Overfitting often occurs when algorithms do not have enough data and learn a very specific configuration. The BDT, once trained is re-applied to the dataset in which the τ_{had} is identified as *tight*. Although the training and the application datasets differ, the BDT is still very performant once applied to the tighter application dataset. The signal samples include both LFV Higgs boson samples produced via ggH, VBF and VH modes. Background samples include all relevant backgrounds from MC simulation and the data-driven misidentified τ_{had} background.

The strategy taken in the BDT approach is to include a mixture of basic four-vector information as well as variables constructed from these four-vectors. While there is a degree of redundancy in this list, so long as the BDT is well trained, this should be no issue. The

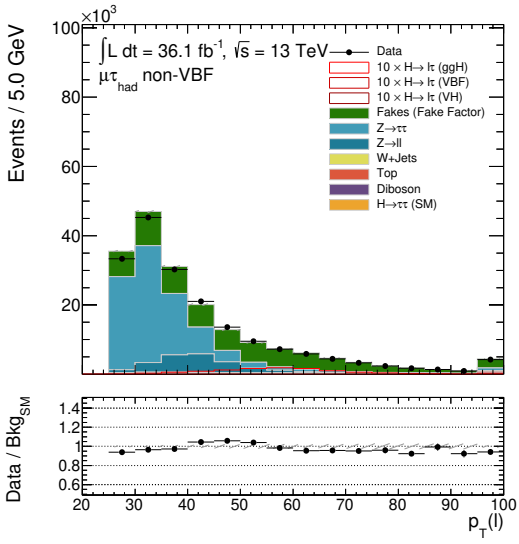
[1] The exact implementation is documented in [144]

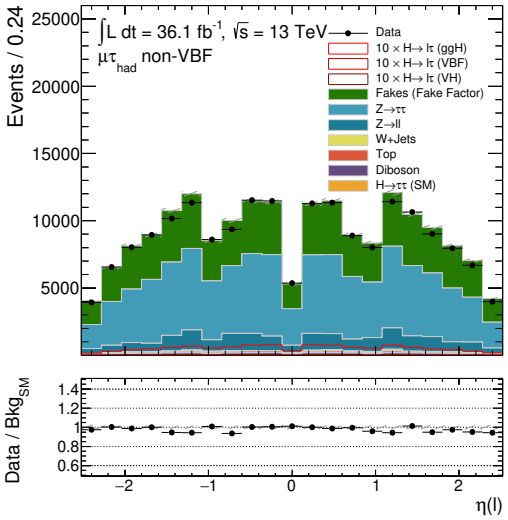
list of inputs are the same for $\mu\tau_{had}$ and $e\tau_{had}$ channels:

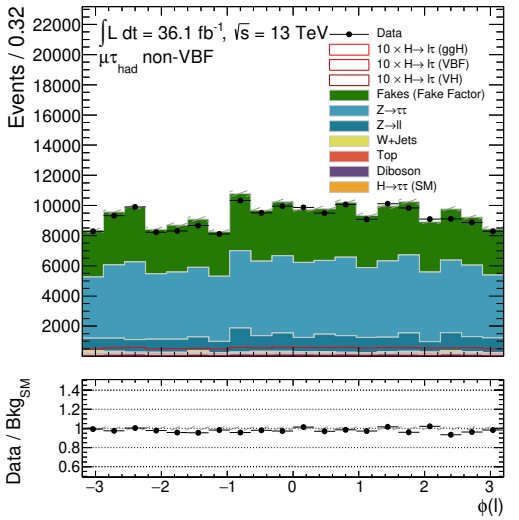
- p_T of the lepton and the τ_{had}. The lepton p_T helps to discriminate prompt leptons (from the Higgs boson decay)
- m_T constructed between either lepton or τ_{had} and E_T^{miss}. The 2D correlation of these variables helps to suppress either W+jets or $Z \to \tau\tau$ events.
- $\Delta\eta(l, \tau_{had})$ and $\Sigma\cos(\Delta\phi)$, two angular variables used to build the preselection definition. These mostly help to reject multijet and W+jet events.
- E_T^{miss}, the missing transverse energy. This is expected be smaller in magnitude in signal compared to ditau backgrounds $Z \to \tau\tau$ and $H \to \tau\tau$ due to the lower number of neutrinos in the signal process.
- m_{vis} and m_{coll}, the visible and collinear masses. The visible mass is particularly useful in separation of $Z \to ee$ backgrounds from the signal. The collinear mass, defined in 5.7, is preferred over the more complex MMC mass as there is only one neutrino, uniquely defined, in this LFV decay. In LFV decays the collinear mass can be expressed as $m_{coll} = \sqrt{2p_T^{\ell}(p_T^{\tau} + E_T^{miss})(\cosh\Delta(\eta) - \cos\Delta(\phi))}$.
- $\Delta R(l, \tau_{had})$, $\eta(l)$, $\eta(\tau)$, $\phi(l)$, $\phi(E_T^{miss})$. Other angular information which can be used in conjunction with other angular information from the lepton, τ_{had} and E_T^{miss} for additional discrimination.

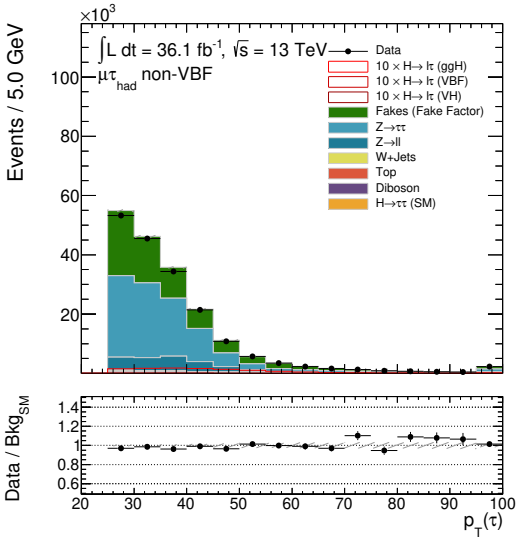
In total, 15 input variables are used for training. The input variables are shown in the non-VBF regions in figures 6.4 and 6.5. Overall the modelling of the variables shows good agreement even given only statistical uncertainties.

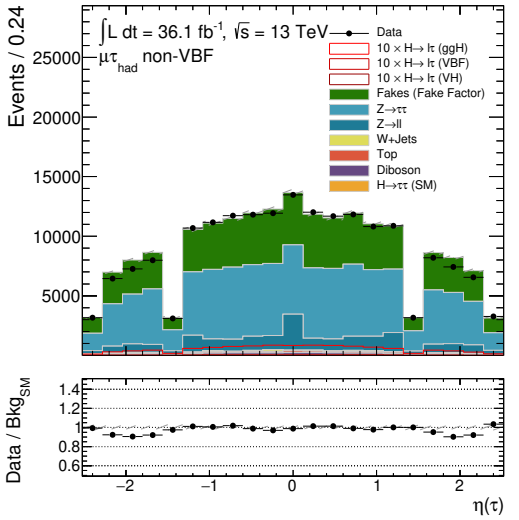
The correlations between inputs variables are presented in figures 6.6 and 6.7. Discerning differences in correlations between background and signal processes are the key to the BDTs effectiveness. Here it can be seen that many correlations exhibit differing behaviour between signal and background. For example, the transverse masses have correlations with each other and with variables such as the visible and collinear masses which differ between signal and background. The lepton p_T in particular has a strong correlation with in background events but not in signal events.

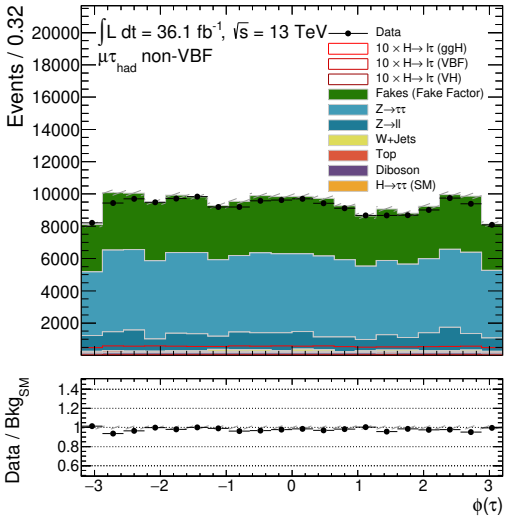

∫L dt = 36.1 fb-1, √s = 13 TeV
μτhad non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Z→ll
W+Jets
Top
Diboson
H→ττ (SM)
Events / 5.0 GeV
Data / BkgSM
pT(l)

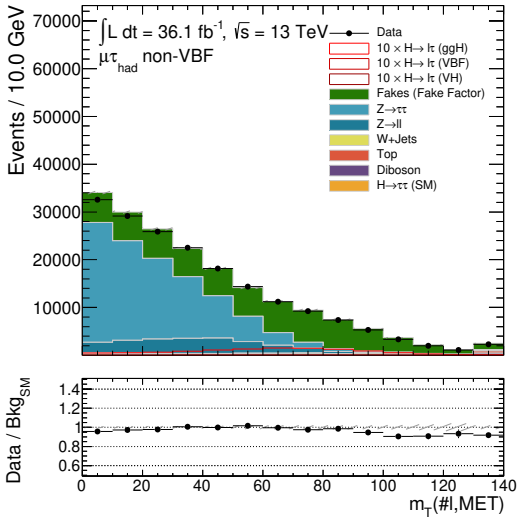

∫L dt = 36.1 fb-1, √s = 13 TeV
μτhad non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Z→ll
W+Jets
Top
Diboson
H→ττ (SM)
Events / 0.24
Data / BkgSM
η(l)

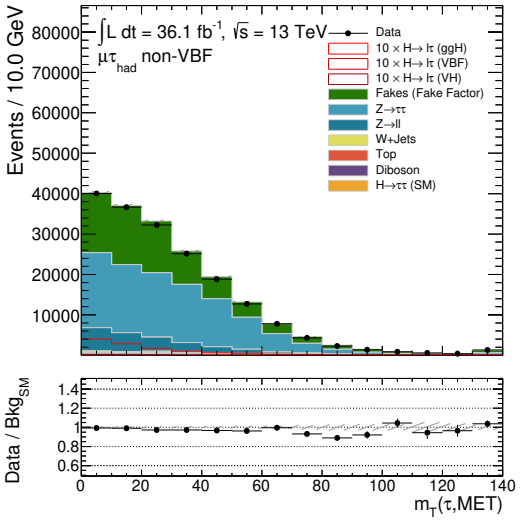

∫L dt = 36.1 fb-1, √s = 13 TeV
μτhad non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Z→ll
W+Jets
Top
Diboson
H→ττ (SM)
Events / 0.32
Data / BkgSM
ϕ(l)

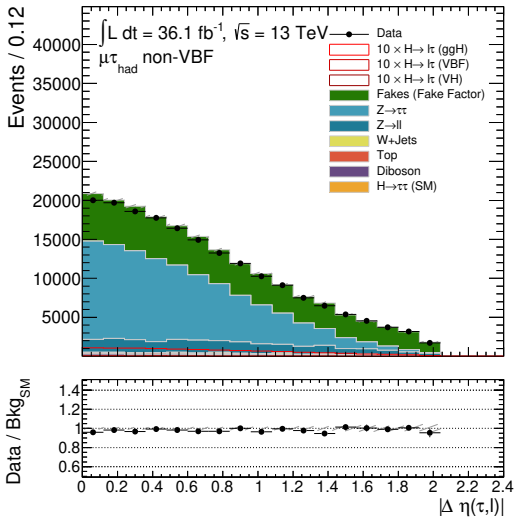

∫L dt = 36.1 fb-1, √s = 13 TeV
μτhad non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Z→ll
W+Jets
Top
Diboson
H→ττ (SM)
Events / 5.0 GeV
Data / BkgSM
pT(τ)

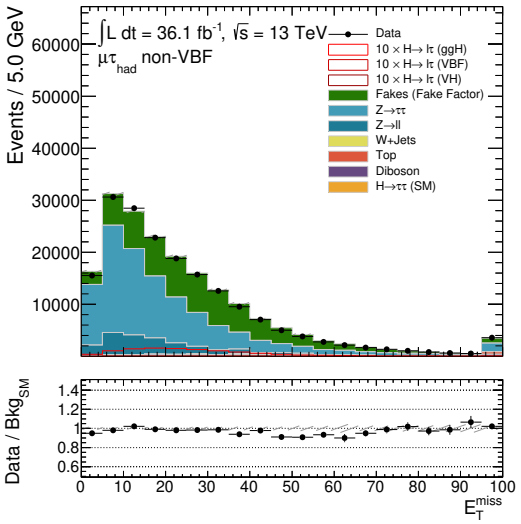

∫L dt = 36.1 fb-1, √s = 13 TeV
μτhad non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Z→ll
W+Jets
Top
Diboson
H→ττ (SM)
Events / 0.24
Data / BkgSM
η(τ)

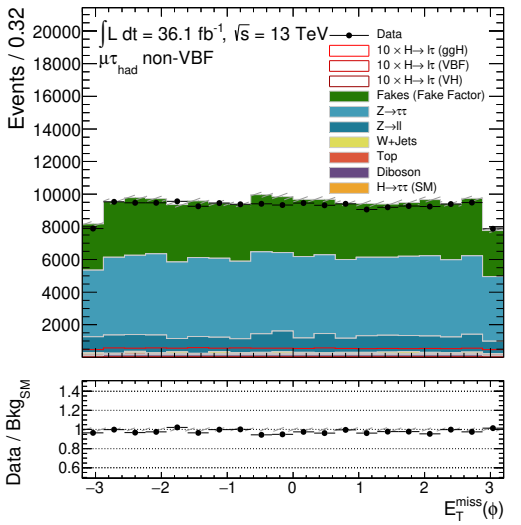

∫L dt = 36.1 fb-1, √s = 13 TeV
μτhad non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Z→ll
W+Jets
Top
Diboson
H→ττ (SM)
Events / 0.32
Data / BkgSM
ϕ(τ)

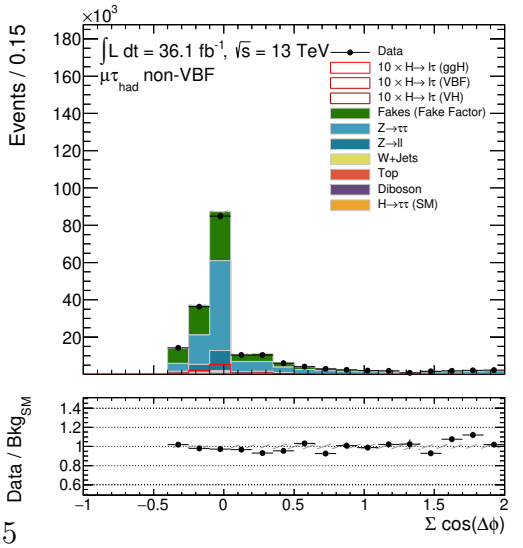

∫L dt = 36.1 fb-1, √s = 13 TeV
μτhad non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Z→ll
W+Jets
Top
Diboson
H→ττ (SM)
Events / 10.0 GeV
Data / Bkg_SM
mT(#l,MET)


∫L dt = 36.1 fb-1, √s = 13 TeV
μτhad non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Z→ll
W+Jets
Top
Diboson
H→ττ (SM)
Events / 10.0 GeV
Data / Bkg_SM
mT(τ,MET)


∫L dt = 36.1 fb-1, √s = 13 TeV
μτhad non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Z→ll
W+Jets
Top
Diboson
H→ττ (SM)
Events / 0.12
Data / Bkg_SM
|Δ η(τ,l)|


∫L dt = 36.1 fb-1, √s = 13 TeV
μτhad non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Z→ll
W+Jets
Top
Diboson
H→ττ (SM)
Events / 5.0 GeV
Data / Bkg_SM
E_T^miss


∫L dt = 36.1 fb-1, √s = 13 TeV
μτhad non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Z→ll
W+Jets
Top
Diboson
H→ττ (SM)
Events / 0.32
Data / Bkg_SM
E_T^miss(ϕ)


×10^3
∫L dt = 36.1 fb-1, √s = 13 TeV
μτhad non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Z→ll
W+Jets
Top
Diboson
H→ττ (SM)
Events / 0.15
Data / Bkg_SM
Σ cos(Δϕ)

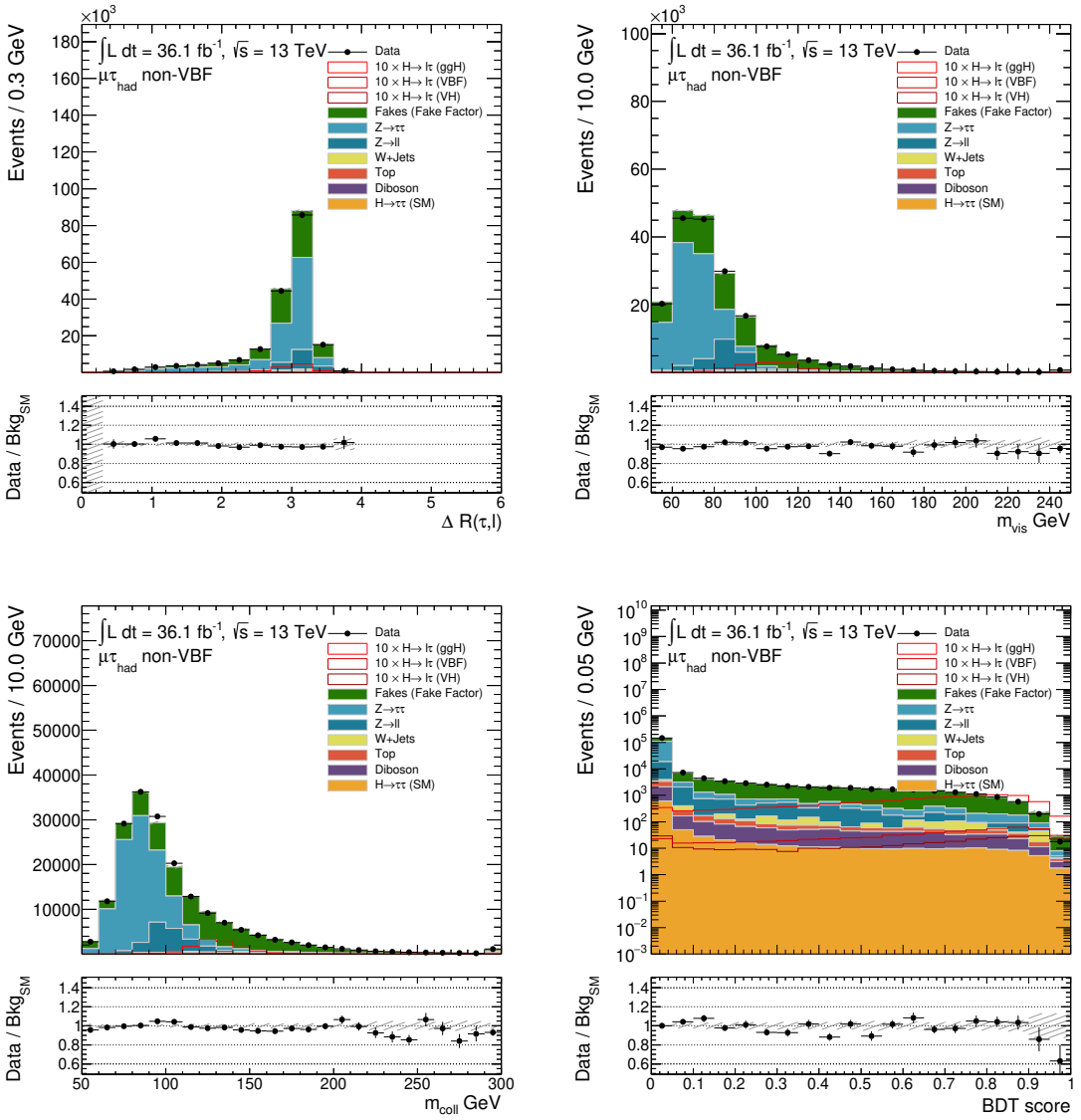
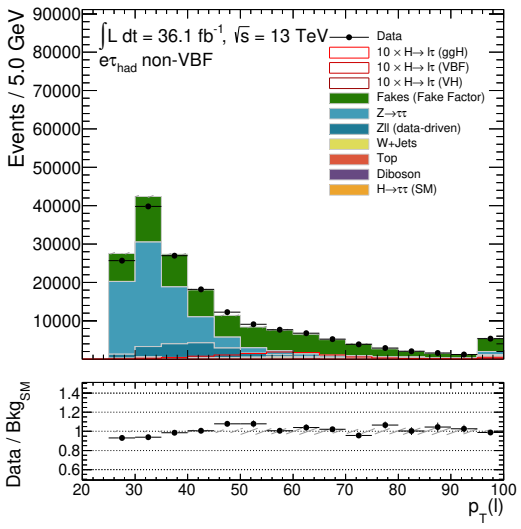
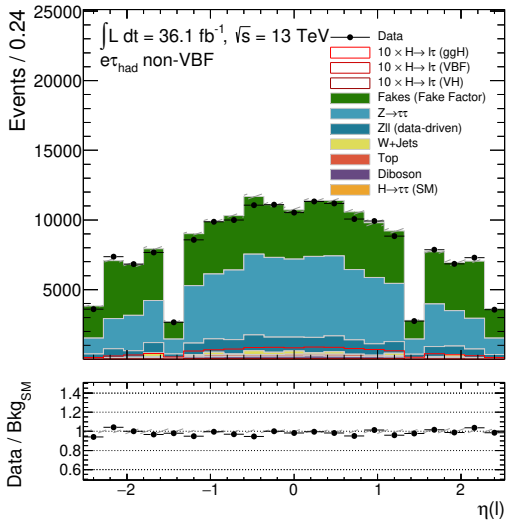
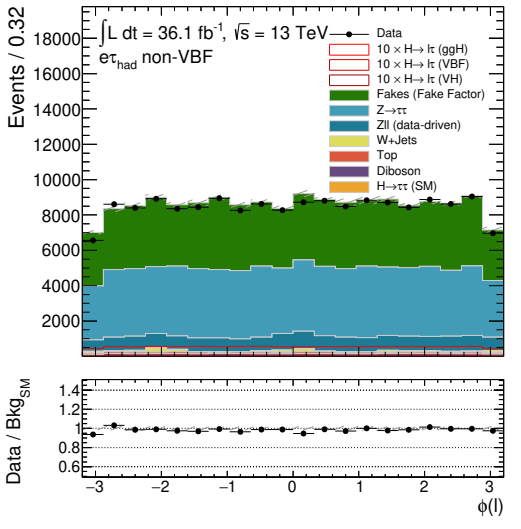
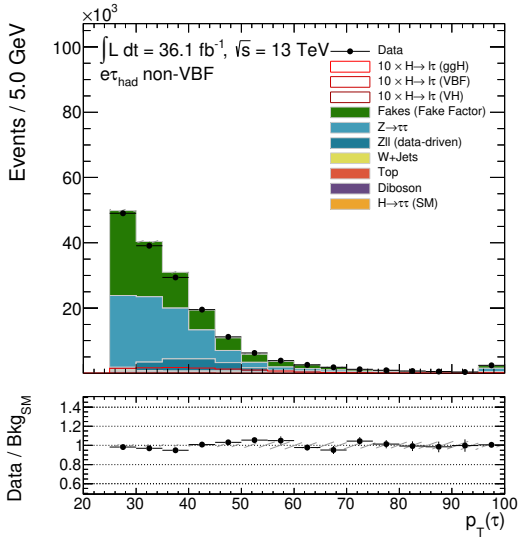


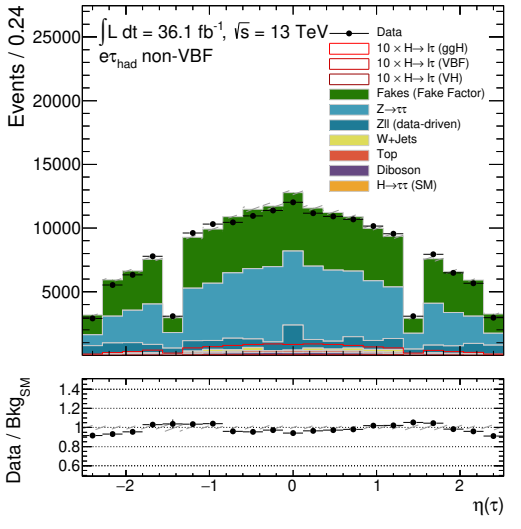
Figure 6.4: Input and output distributions for the BDT in the $\mu\tau_{had}$ non-VBF region. The first two rows contain the p_T, η and ϕ distributions (left to right) for the lepton then the τ_{had} (top to bottom) respectively. The third row contains the m_T distributions between the lepton and E_T^{miss}, then the τ_{had} and E_T^{miss} as well as the $\Delta\eta$ between lepton and τ_{had} (left to right) respectively. The fourth row contains the E_T^{miss}, its direction in the transverse plane and the $\Sigma\cos(\Delta\phi)$ distribution. The final row contains the ΔR between lepton and τ_{had} candidates as well as the visible then the collinear approximated masses (left to right) respectively. Uncertainty bands are only statistical uncertainties.

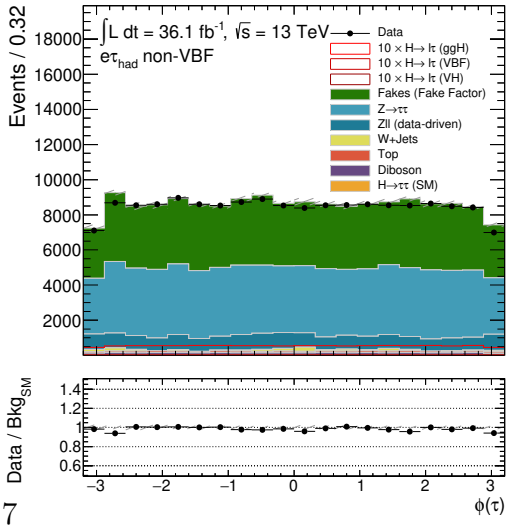

∫L dt = 36.1 fb^{-1}, √s = 13 TeV
eτ$_{had}$ non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Zll (data-driven)
W+Jets
Top
Diboson
H→ττ (SM)
Events / 5.0 GeV
Data / Bkg$_{SM}$
p$_T$(l)

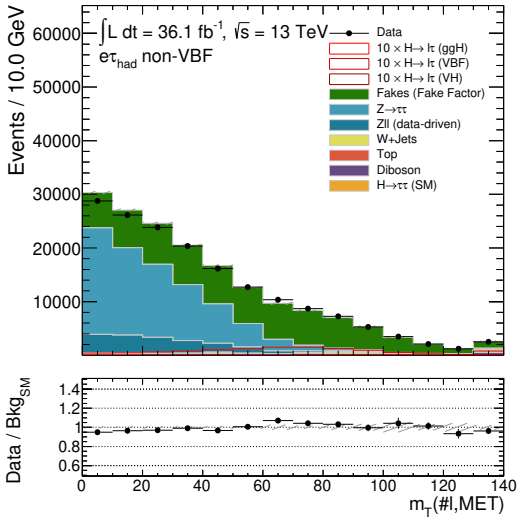

∫L dt = 36.1 fb^{-1}, √s = 13 TeV
eτ$_{had}$ non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Zll (data-driven)
W+Jets
Top
Diboson
H→ττ (SM)
Events / 0.24
Data / Bkg$_{SM}$
η(l)

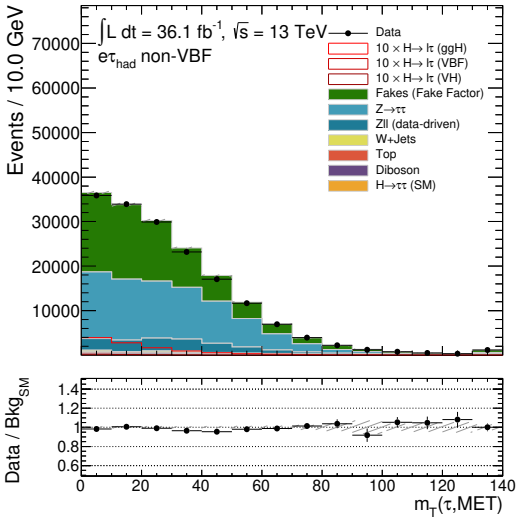

∫L dt = 36.1 fb^{-1}, √s = 13 TeV
eτ$_{had}$ non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Zll (data-driven)
W+Jets
Top
Diboson
H→ττ (SM)
Events / 0.32
Data / Bkg$_{SM}$
ϕ(l)

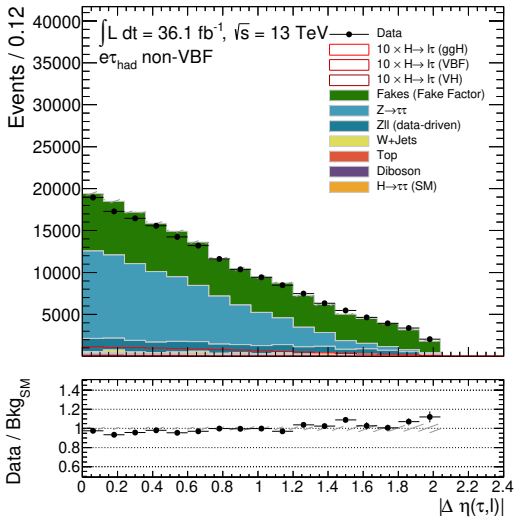

×10^3
∫L dt = 36.1 fb^{-1}, √s = 13 TeV
eτ$_{had}$ non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Zll (data-driven)
W+Jets
Top
Diboson
H→ττ (SM)
Events / 5.0 GeV
Data / Bkg$_{SM}$
p$_T$(τ)

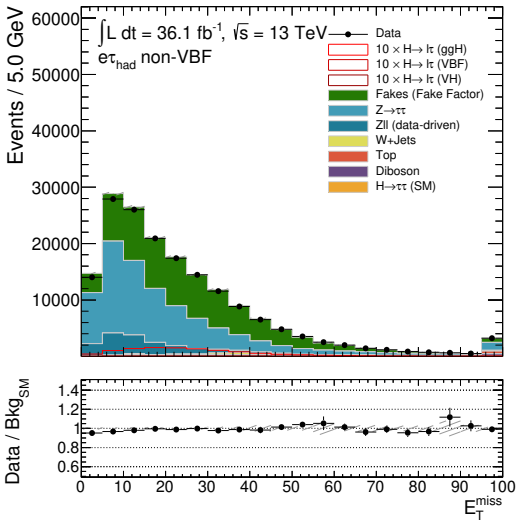

∫L dt = 36.1 fb^{-1}, √s = 13 TeV
eτ$_{had}$ non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Zll (data-driven)
W+Jets
Top
Diboson
H→ττ (SM)
Events / 0.24
Data / Bkg$_{SM}$
η(τ)

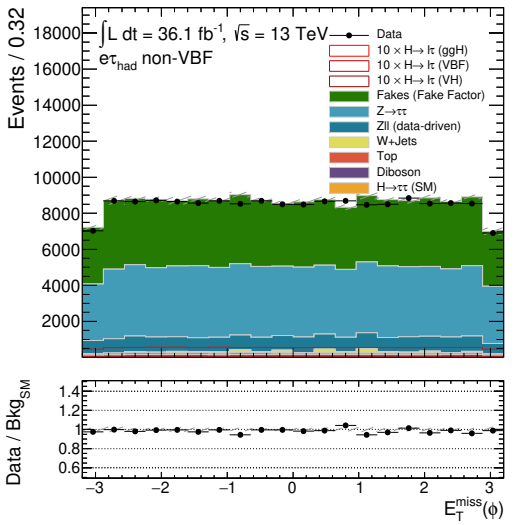

∫L dt = 36.1 fb^{-1}, √s = 13 TeV
eτ$_{had}$ non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Zll (data-driven)
W+Jets
Top
Diboson
H→ττ (SM)
Events / 0.32
Data / Bkg$_{SM}$
ϕ(τ)

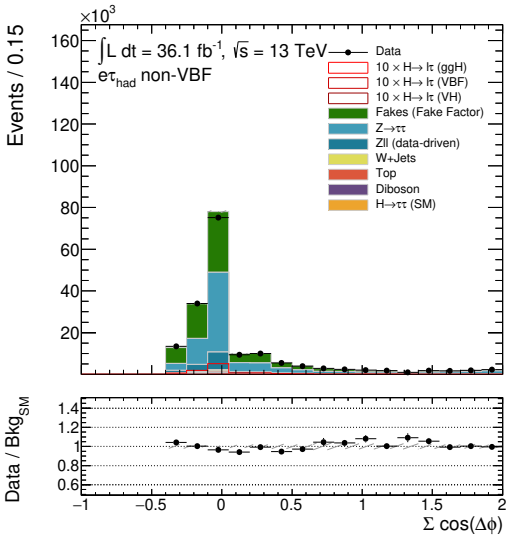

∫L dt = 36.1 fb^{-1}, √s = 13 TeV
eτ$_{had}$ non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Zll (data-driven)
W+Jets
Top
Diboson
H→ττ (SM)
Events / 10.0 GeV
Data / Bkg$_{SM}$
m$_T$(#l,MET)


∫L dt = 36.1 fb^{-1}, √s = 13 TeV
eτ$_{had}$ non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Zll (data-driven)
W+Jets
Top
Diboson
H→ττ (SM)
Events / 10.0 GeV
Data / Bkg$_{SM}$
m$_T$(τ,MET)


∫L dt = 36.1 fb^{-1}, √s = 13 TeV
eτ$_{had}$ non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Zll (data-driven)
W+Jets
Top
Diboson
H→ττ (SM)
Events / 0.12
Data / Bkg$_{SM}$
|Δ η(τ,l)|


∫L dt = 36.1 fb^{-1}, √s = 13 TeV
eτ$_{had}$ non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Zll (data-driven)
W+Jets
Top
Diboson
H→ττ (SM)
Events / 5.0 GeV
Data / Bkg$_{SM}$
E$_T^{miss}$


∫L dt = 36.1 fb^{-1}, √s = 13 TeV
eτ$_{had}$ non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Zll (data-driven)
W+Jets
Top
Diboson
H→ττ (SM)
Events / 0.32
Data / Bkg$_{SM}$
E$_T^{miss}$(ϕ)


×10^3
∫L dt = 36.1 fb^{-1}, √s = 13 TeV
eτ$_{had}$ non-VBF
Data
10 × H→ lτ (ggH)
10 × H→ lτ (VBF)
10 × H→ lτ (VH)
Fakes (Fake Factor)
Z→ττ
Zll (data-driven)
W+Jets
Top
Diboson
H→ττ (SM)
Events / 0.15
Data / Bkg$_{SM}$
Σ cos(Δϕ)

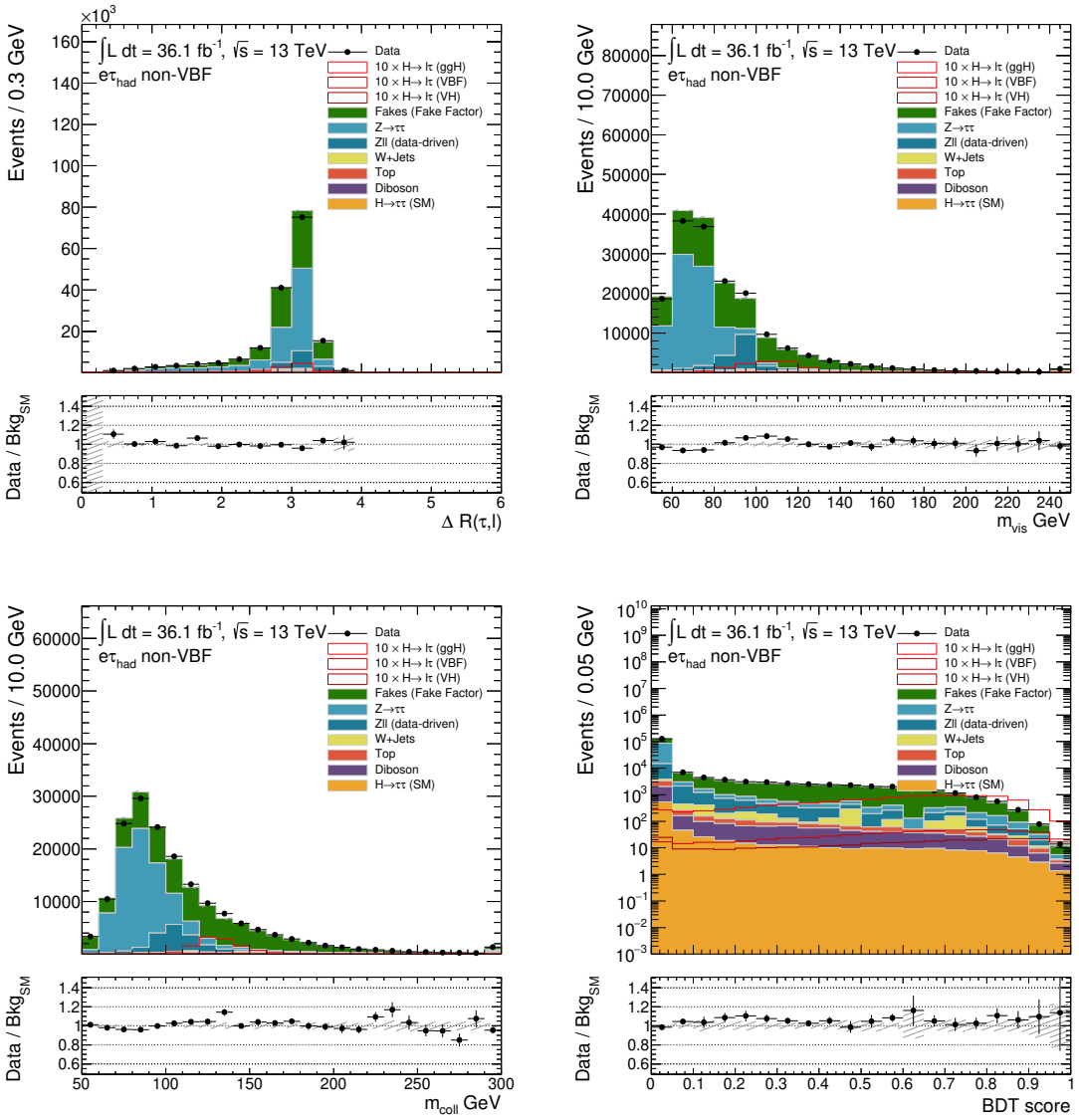


Figure 6.5: Input and output distributions for the BDT in the $e\tau_{had}$ non-VBF region. The first two rows contain the p_T, η and ϕ distributions (left to right) for the lepton then the τ_{had} (top to bottom) respectively. The third row contains the m_T distributions between the lepton and E_T^{miss}, then the τ_{had} and E_T^{miss} as well as the $\Delta\eta$ between lepton and τ_{had} (left to right) respectively. The fourth row contains the E_T^{miss}, its direction in the transverse plane and the $\Sigma\cos(\Delta\phi)$ distribution. The final row contains the ΔR between lepton and τ_{had} candidates as well as the visible then the collinear approximated masses (left to right) respectively. Uncertainty bands are only statistical uncertainties.

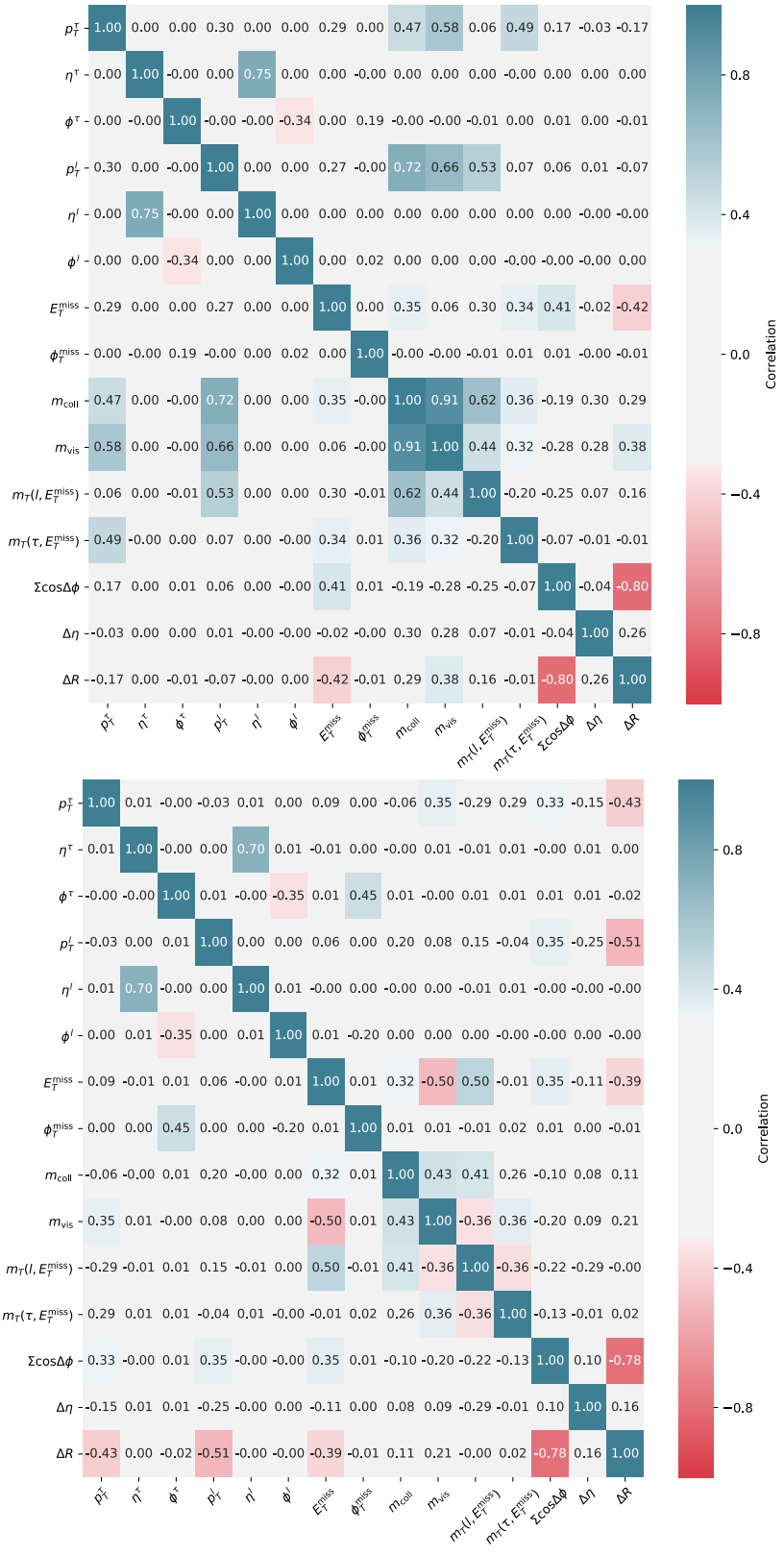


Figure 6.6: The correlations between input variables for background (top) and signal (bottom) processes for the $H \to e\tau$ search.

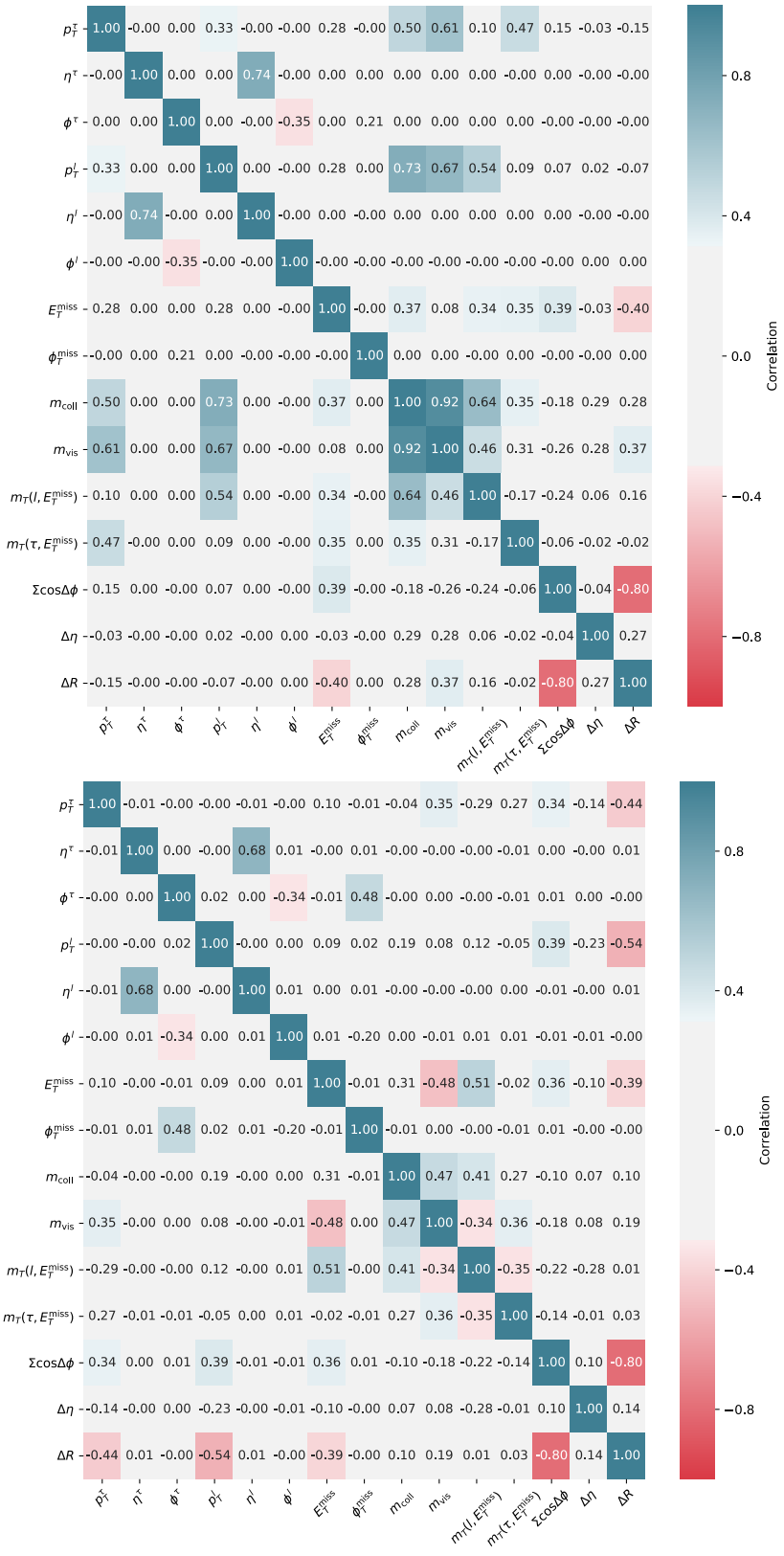


Figure 6.7: The correlations between input variables for background (top) and signal (bottom) processes for the $H \to \mu\tau$ search.

6.4.3 Training Strategy

The training sample is initially separated into two subsamples (per channel) based on the "parity" of the event (which is based on the event number). Events (both signal and background) with an even event number are assigned to the "even" subsample and those with odd event number are assigned to the "odd" subsample. For each search, two BDTs are trained independently either even or odd datasets and then applied to the opposing parity dataset. This procedure is done to avoid overtraining a BDT which is often the case when the training and application dataset are not statistically independent. Events in each subsample are then further subdivided into a training, validation and test dataset, consisting of 80%, 10% and 10% of the subsample (randomly assigned). Each dataset has a specific purpose which will be described below.

The *training* dataset is utilised to train a given BDT with a specific configuration of hyperparameters, which dictate the structure/complexity of the BDT model. The BDTs performance with a specific configuration of hyperparameters is calculated by the AUC score calculated on the *validation* dataset (which is independent of the *training* dataset). The AUC scores calculated on the *validation* dataset are then used to select the BDT with optimal hyperparameters in a procedure described next. The *test* dataset is used to compare the performance of a hold-out dataset of the optimal BDT. Overtraining would reveal a weaker performance on the *test* dataset. For each BDT, the training dataset consists of approximately 800000 background and 20000 signal raw events.

A hyperparameter scan is performed using Bayesian Optimisation scan [145]. As the AUC score is a function of the hyperparameters, which may not be a smooth differentiable function, one must employ more sophisticated methods for maximising the AUC score. The Bayesian method used begins by guessing a set of random hyperparameter values. A prior function is built to model the behaviour of the AUC with respect to the hyperparameters. After each iteration of the scan, the prior is updated based on information received from the scan (namely the AUC score) and the posterior distribution is then used to determine the next point to evaluate. This is repeated for 200 iterations after a random sampling of 50 guesses.

The hyperparameters scanned over for the BDT optimisation are:

- "Max depth": the maximum depth on any given decision tree
- "Number of trees": the number of boosts performed in the ensemble
- "Learning rate (also known the shrinkage)": how quickly the error (misclassification) reduces with respect to previous tree

The fine tuning of hyperparameters can lead to overtrained BDTs so a check must be performed in order to assess this. The most performant MVA, based on the AUC score calculated on the validation dataset, is tested against the performance on the test dataset to

assess overtraining. No such overtraining was found. As both validation and testing datasets are independent of the training data, the MVA should be completely unbiased with respect to the training data. Once an optimal MVA is selected, this MVA is applied to the opposite parity events. The schematic displayed in figure 6.8 summarises this procedure. The optimal hyperparameter values for depth vary between 3 and 10 with the number of trees being between 700 and 1000. Optimal learning rates were determined to be between 0.20 and 0.09.

A measure of the strength of certain input features is the "importance" of a variable

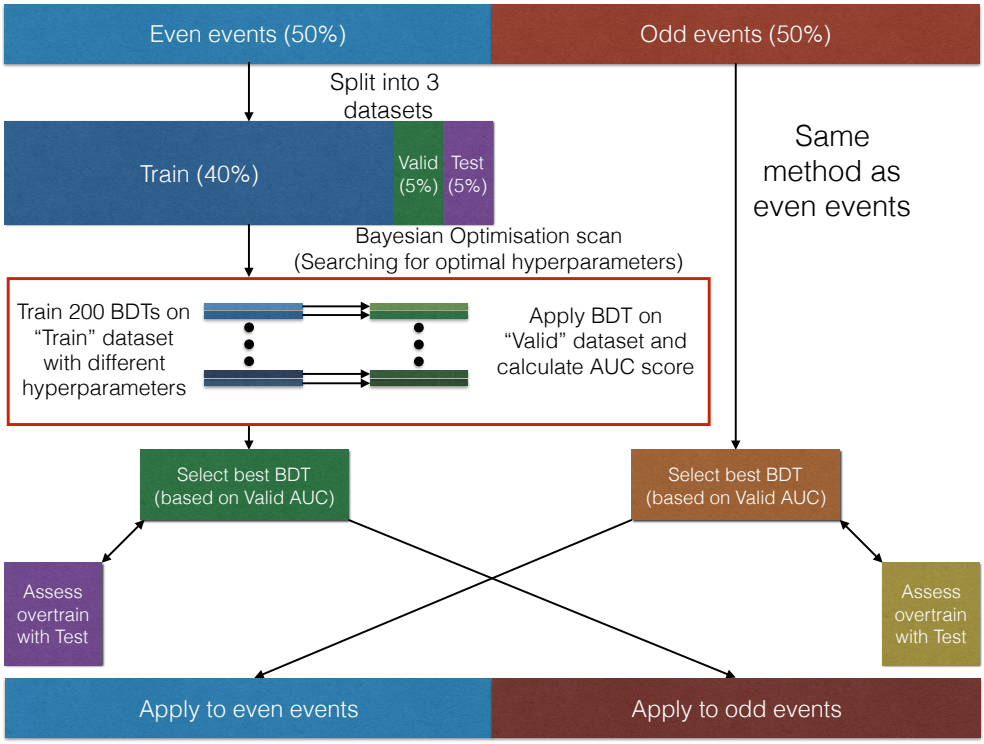


Figure 6.8: A schematic of the BDT training, optimisation and evaluation procedure.

[146]. The importance is calculated by performing a weighted sum over nodes in the tree, assessing the performance of how a particular input feature splits a node. This is a measure of how effective a feature performs in classifying events. The measure taken to assess such effectiveness is the Gini Index which is commonly used to measure the disparity between two populations in two samples. Larger scores correspond to more important features in distinguishing the signal events from background. The importances for the input features for each of the BDTs are presented in 6.9.

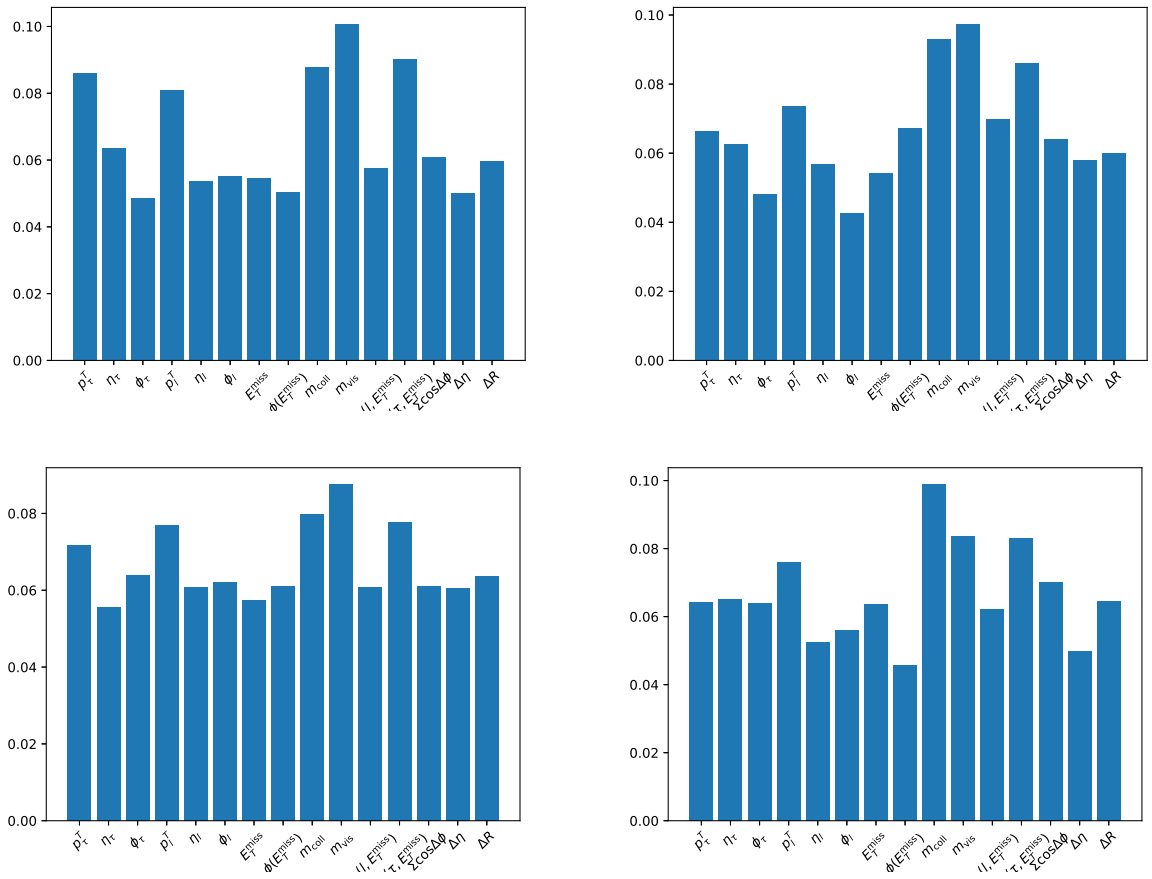


Figure 6.9: Importance of input features for the BDT in the $e\tau$ ($\mu\tau$) channels on top (bottom) rows trained on even (odd) parity events on the left (right).

Generally speaking the most important inputs are the mass variables, the p_T of the lepton and τ_{had} as well as the m_T variables. These reflect the findings from correlations between variables which were discussed in section 6.4.2.

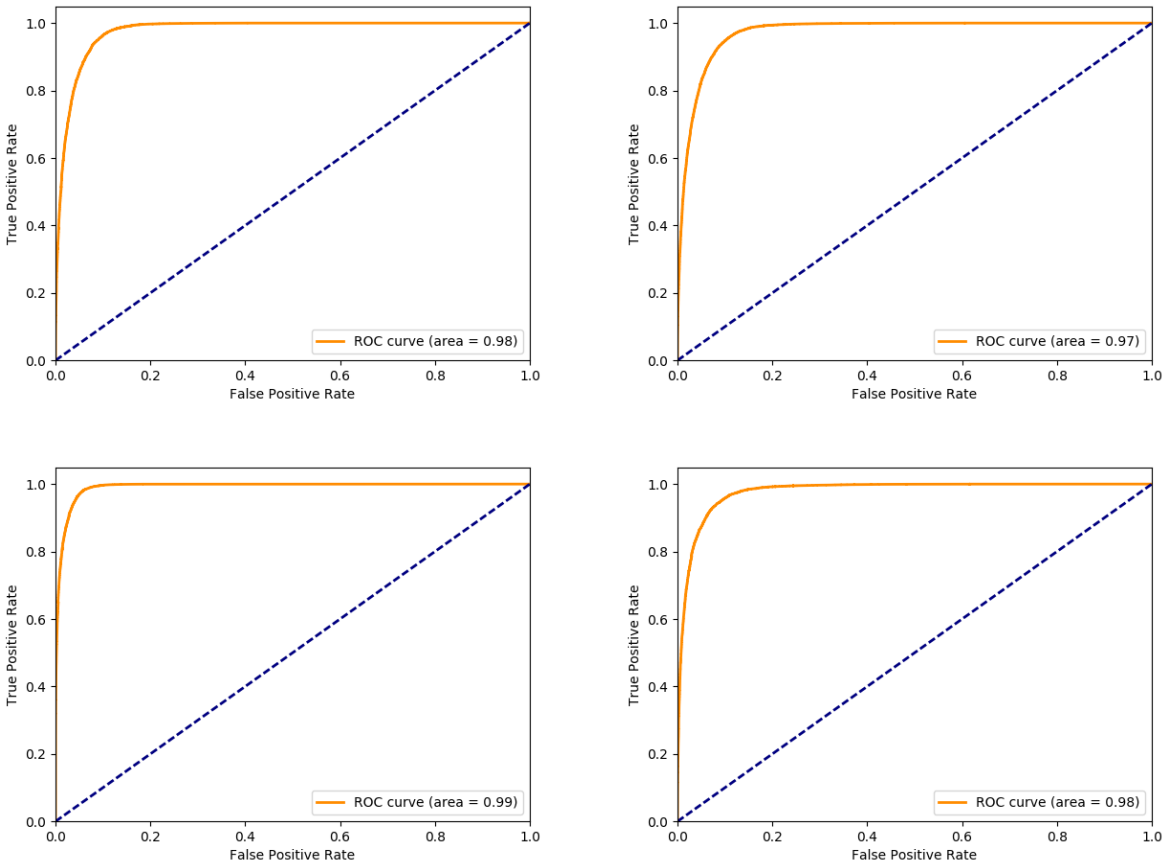


Figure 6.10: ROC curves for the test dataset for BDTs trained even (odd) parity events seen left (right). Plots are presented top (bottom) for the $e\tau_{had}$ and $\mu\tau_{had}$ channels.

The performance of the BDT is summarised in the plots 6.10. These show good performance overall, with an area under the ROC curve of 0.98.

6.5 Background Estimation

The background composition are illustrated in 6.2. The dominant background sources are from $Z \to \tau\tau$ and events where the τ_{had} a misidentified jet or lepton.

The $Z \to \tau\tau$ background, constituting 44-50% of the background, is modelled with a MC based estimate with a normalisation calculated by the fit. The normalisation is correlated to a $Z \to \tau\tau$ control region which was defined in section 6.3.

Events with jets misidentified as a τ_{had} contributes 35-40% of the background. These backgrounds are modelled with the data-driven fake-factor method which was defined in section 5.6.

In the $H \to e\tau$ search a subleading background is due to events where *electrons* are misidentified as τ_{had}, primarily from $Z \to ee$ events. This background contributes approximately 10% of the background. A data-driven technique, similar to the fake factor, was developed to model with these backgrounds.

All other backgrounds, each contributing less than than 10% to total background, is modelled with MC simulation.

6.5.1 Backgrounds from Jets Misidentified as τ_{had}

The method for modelling backgrounds sources where jets fake the τ_{had} signature is detailed in section 5.6. The only differences accounted for are due to the contrasting phase spaces.

Fake factors are measured in control regions which are defined in section 6.3 to be orthogonal to the signal regions. The fake factors are calculated as the ratio of events with an identified τ_{had} to events with an identified anti-τ_{had}. The τ_{had} candidates are required to pass the *tight* working point. Conversely the anti-τ_{had} candidate is required pass a BDT score of 0.5 but fail the *tight* working point. The weighted average between fake factors for W+jets and multijet events is applied to signal region events requiring a selected anti-τ_{had} rather than a selected τ_{had}.

The leading uncertainties associated with the estimation are due to the statistical limitation of the fake factor calculation as well as the non-closure of the method. The uncertainty due to the calculation of the fake factor contributes to a 16% difference in the yield of the background with respect to the nominal estimate. The non-closure (calculated by reversing the roles of OS and SS regions) contributes a significant shape variation of the misidentified τ_{had} background in the BDT classifier score distribution.

6.5.2 Backgrounds from Electron Misidentified as τ_{had}

The electron signature can also very closely mimic the one-pronged τ_{had} signature[2]. In particular the $Z \to ee$ background produces a large number of electrons which can be misidentified as a τ_{had} candidate. A data-driven estimation method has been used to model this background.

The data-driven method involves the use of a fake factor for electrons misidentified as τ_{had}which is similar to the fake factor used for modelling jets misidentified as τ_{had}. The electron fake factor measures the rate at which an electron passes τ_{had} identification to the rate at which the electron fails the τ_{had} identification. The expected number of events from $Z \to ee$ events (only in one pronged events) is estimated as:

$$N_{e\to\tau_{had}} = FF_{e\to\tau_{had}} \times N_{\text{SR, fail electron veto}}, \tag{6.1}$$

where $FF_{e\to\tau_{had}}$ is a data-driven fake-factor and the $N_{\text{SR, fail electron veto}}$ are the number of events failing the electron veto described in section 4.2. Note events with a real τ_{had} there are implicitly subtracted here. Also subtracted are the cases where jets are misidentified, modelled using the standard fake factor method.

The determination of this $FF_{e\to\tau_{had}}$ is calculated in a control region in which the $Z \to ee$ is enriched. This region requires a τ_{had} identification requirement passing the *medium but not tight* working point A 5 GeV mass window in m_{vis} about the Z mass in conjunction with a low transverse mass ($m_T(l, E_T^{miss}) < 40$ GeV and $m_T(\tau_{had}, E_T^{miss}) < 60$ GeV) is used to define the $Z \to ee$ control region. The ratio calculated in this fake factor measures the rate at which events pass and fail the electron-τ_{had} veto BDT requirement. The ratio also contains the subtraction of events which are not electrons misidentified as τ_{had}.

Systematic uncertainties are calculated in the same manner as the fake factor for jets misidentified. An additional uncertainty is added to account for the difference in fake factor calculated in a region where τ_{had} are identified with a *tight* as opposed to identified as *medium but not tight*. Of these uncertainties, the leading systematic effect is due to this difference in fake factors calculated with different identification working point, producing an 18% variation in the yield.

[2]The three-pronged τ_{had} are modelled with MC as electrons rarely mimic this signature

6.6 Fit Model and Systematics

Identical procedure to section 5.9.3 is implemented set up the fit model. This includes experimental systematics described in section 5.8. Uncertainties relating to the data-driven modelling of fake backgrounds were described in section 6.5.

The $Z \to \tau\tau$ theory systematics are treated with the method described in section 5.9.3. Theory systematics are correlated across channels and categories (VBF and the non-VBF) in the shape component only. The normalisation component is derived from two floating normalisations in the fit, which correlated between signal regions and $Z \to \tau\tau$ control regions. One component normalises the background in the non-VBF channel and the other for VBF.

The signal systematics, which were described in section 5.8.2, are applied here for both the SM $H \to \tau\tau$ background and the LFV signal. Systematics are treated as correlated between the SM and LFV Higgs boson decays.

The input to the fit model was either the $H \to e\tau$ or $H \to \mu\tau$ signals with a branching ratio set to 1%. The parameter of interest is the branching ratio of the $H \to e\tau$ or $H \to \mu\tau$. This parameter is initially set to 1% as a benchmark test. The SM $H \to \tau\tau$ cross-section is set to the SM expectation. Note that the BR of $H \to e\tau$ and $H \to \mu\tau$ are treated as completely uncorrelated as two fits are performed (one for $H \to e\tau$ and one for $H \to \mu\tau$).

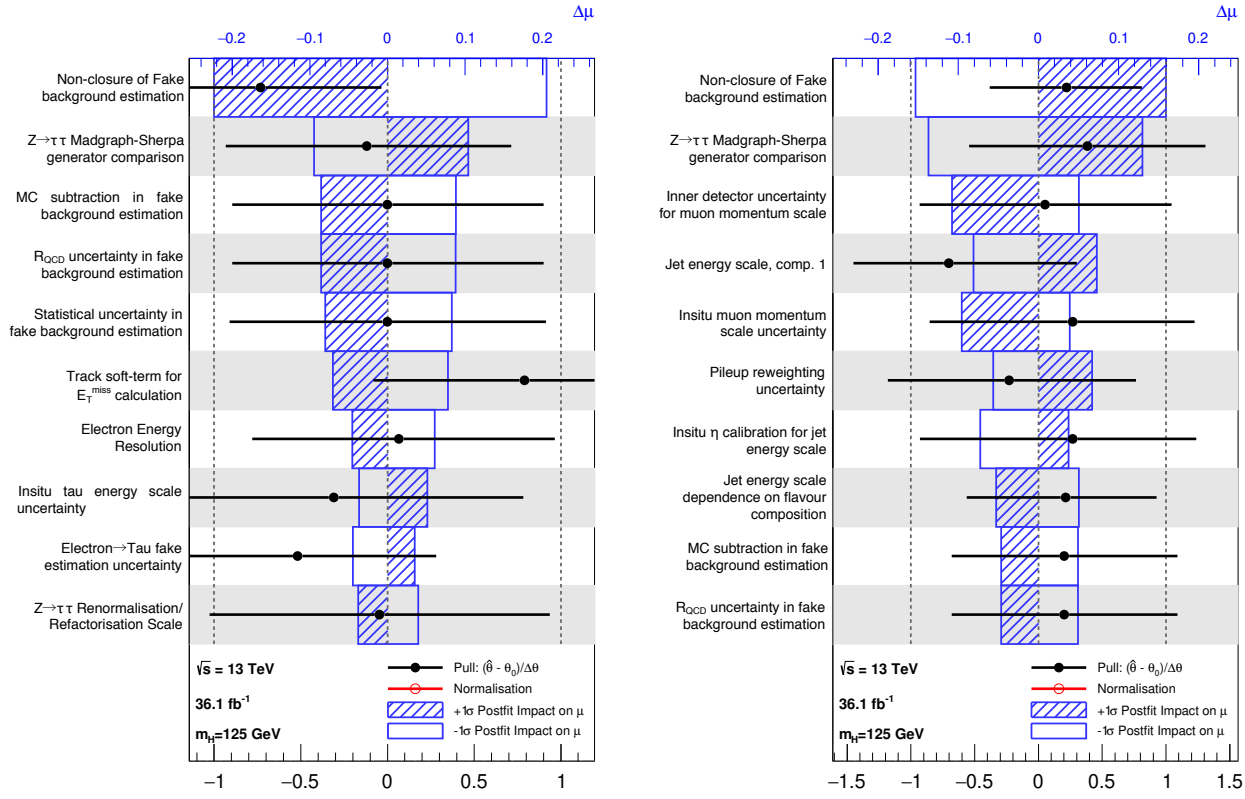
The impacts are summarised in table 6.2. Overall, the measurement is rather systematically limited. The largest systematics originate from the misidentified τ_{had} backgrounds and from the jet calibration uncertainties. The large uncertainty in the background modelling indicates a need to improve the background estimates in future studies. The suppression of $Z \to \tau\tau$ background leaves the fake backgrounds contributing in the signal sensitive bins of the BDT distribution. The impact due to the reconstruction and calibration of the leptons have large role in this search as much of the sensitivity derives from the p_T spectra of the lepton.

Source of uncertainty	Impact on $\Delta\mu(H \to e\tau)$			Impact on $\Delta\mu(H \to \mu\tau)$		
	Measured	Expected	Rank	Measured	Expected	Rank
Electron	+0.05/-0.05	+0.07/-0.07	3	+0.03/-0.03	+0.04/-0.03	10
Muon	+0.04/-0.04	+0.04/-0.04	5	+0.10/-0.10	+0.09/-0.09	3
Tau lepton	+0.02/-0.02	+0.03/-0.03	9	+0.04/-0.04	+0.06/-0.05	8
Jets	+0.09/-0.08	+0.08/-0.08	2	+0.11/-0.12	+0.13/-0.12	1
$E_{\mathrm{T}}^{\mathrm{miss}}$	+0.02/-0.02	+0.04/-0.03	10	+0.05/-0.08	+0.03/-0.03	5
b-tag	+0.02/-0.03	+0.04/-0.03	8	+0.01/-0.01	+0.02/-0.01	12
Misidentified τ_{had}	+0.13/-0.12	+0.13/-0.13	1	+0.11/-0.11	+0.13/-0.12	2
Pileup reweighting	+0.02/-0.01	+0.04/-0.02	11	+0.05/-0.03	+0.10/-0.08	7
Overall Ztt	+0.01/-0.01	+0.02/-0.02	13	+0.00/-0.00	+0.02/-0.01	14
Luminosity	+0.01/-0.00	+0.03/-0.02	14	+0.01/-0.00	+0.03/-0.02	13
Background norm.	+0.05/-0.04	+0.08/-0.06	4	+0.04/-0.02	+0.11/-0.08	11
Theor. uncert. (bkg)	+0.04/-0.03	+0.05/-0.04	7	+0.08/-0.07	+0.10/-0.09	4
Theor. uncert. (sig)	+0.01/-0.01	+0.06/-0.04	12	+0.04/-0.02	+0.07/-0.05	9
MC statistics	+0.04/-0.04	+0.04/-0.04	6	+0.04/-0.04	+0.06/-0.06	6
Full Syst.	+0.17/-0.16	+0.20/-0.19		+0.18/-0.18	+0.24/-0.22	
Data Stat.	+0.07/-0.07	+0.08/-0.08		+0.07/-0.07	+0.09/-0.08	
Total	+0.18/-0.17	+0.21/-0.20		+0.19/-0.19	+0.25/-0.24	

Table 6.2: Expected and observed impact on the $\Delta\mu$ by various sources of uncertainty. Both expected results are calculated with respect to a LFV signal corresponding to a BR of 1%.

6.7 Results

The fits for both non-VBF hadronic decay channels were stable with the rankings of high impacting nuisance parameters shown in figure 6.11. No pulls exceed 1 σ and none are overly constrained other than fake uncertainties. This is the case as the non-VBF region is very inclusive so has high statistical power to constrain the conservative uncertainties for the fake background estimations. The high ranking nuisance parameters were discussed in the previous section.



(a) $e\tau_{had}$ non-VBF channel

(b) $\mu\tau_{had}$ non-VBF channel

Figure 6.11: The top ten nuisance parameters ranked by their impact on μ (blue). The black markers show the pulls of the nuisance parameters.

The output distribution for the non-VBF channel is presented in figure 6.12 and post-fit yields are tabulated in table 6.3. A slight deficit in data, compared to the expected background, is seen in the $H \to \mu\tau$ search. Compatibility with respect to the background expectation is observed in the $H \to e\tau$ channel. No significant excess was observed in combination with

other channels (VBF and fully leptonic decays). The best-fit branching ratio measured is $0.15^{+0.19}_{-0.17}\%$ and $-0.22^{+0.19}_{-0.19}\%$ for the $H \to e\tau$ and $H \to \mu\tau$ searches, respectively. Both results are consistent with no presence of LFV Higgs boson decays.

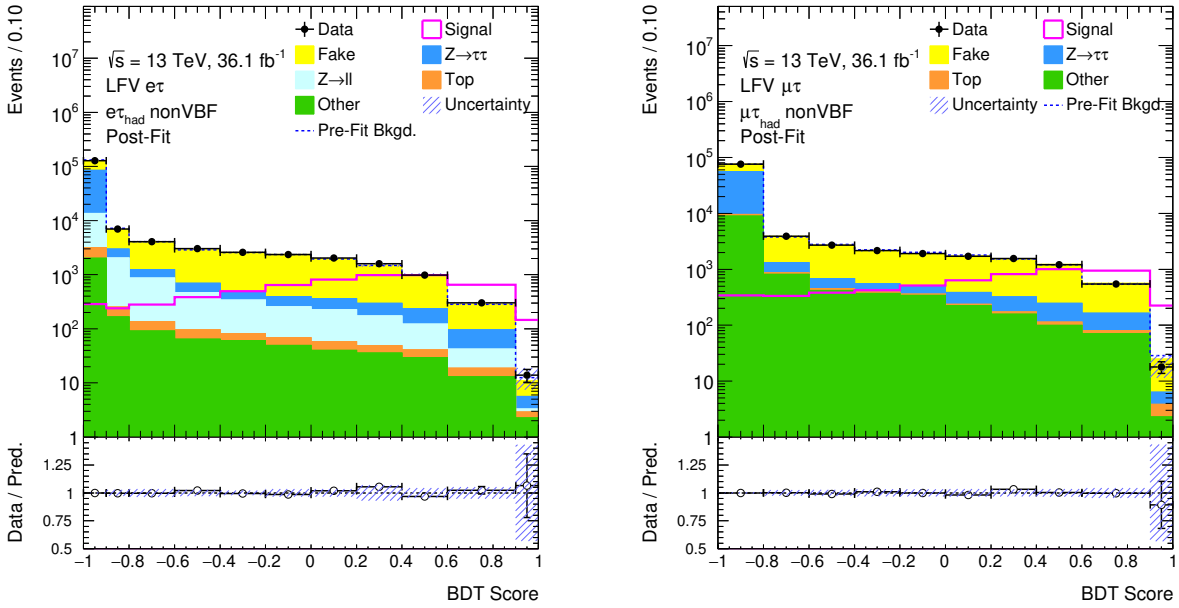


Figure 6.12: The distribution of the BDT after the fit. The $H \to e\tau$ search is shown on the left and $H \to \mu\tau$ on the right. The lower panel represents the data/predicted background ratio. Note that the signal is presented, unstacked, with a BR of 1% in both plots.

	$\mu\tau_{\text{had}}$	$e\tau_{\text{had}}$
$Z \to \tau\tau$	95800 ± 1800	74400 ± 2100
Top	1620 ± 210	1590 ± 190
Fake	64400 ± 1600	73500 ± 1900
$Z \to \ell\ell$	-	16000 ± 1800
Other	22900 ± 1000	2951.4 ± 200
Total Bkg.	185000.0 ± 1000	168700.0 ± 950
Signal	-260 ± 250	170 ± 220
Data	184508	168883

Table 6.3: Event yields in the non-VBF hadronic channel. Uncertainties include statistical and systematic uncertainties. "Other" are the minor background processes, such as diboson and $H \to \tau\tau$. In the $H \to \mu\tau$ search $Z \to \ell\ell$ is also included on "Other". Signal event yields reflect the best-fit branching ratios stated in text.

Upper limits at the 95% confidence level are placed on the branching ratio of potential $H \to e\tau$ and $H \to \mu\tau$ decays, assuming a SM cross-section for the production of the Higgs boson. The summary of the breakdown of upper limits per search, channel and category are presented in figure 6.13. The sensitivity is largely driven by the non-VBF categories as the large signal acceptance combined with the good background discrimination, provided by the BDT, allows for an optimal sensitivity.

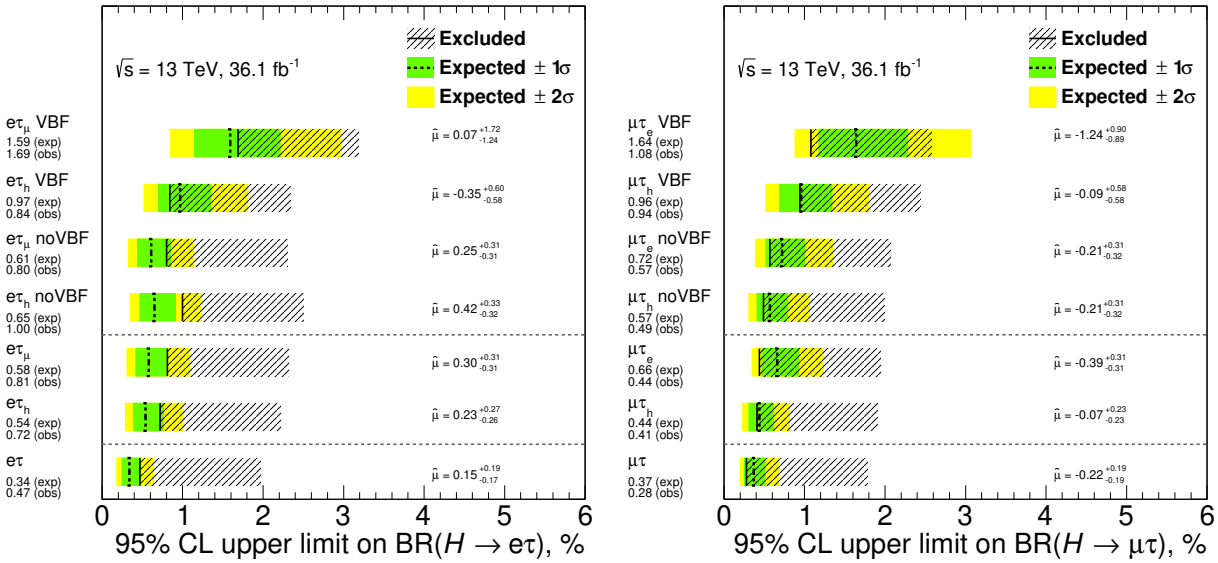


Figure 6.13: The upper limit of $H \to e\tau$ and $H \to \mu\tau$ branching ratios on the left and right plots respectively. The best fit BR is also stated on the right hand side of the plot. The combinations in each of the channels as well as a full combination are presented in the bottom two sections of the plot respectively. The uncertainty band contains statistical and systematic uncertainties.

The observed upper limits are 0.47% and 0.28% for the $H \to e\tau$ and $H \to \mu\tau$ decays respectively, driven largely by the non-VBF categories (particularly in the hadronic channel). These limits are competitive with the CMS upper limits limits of 0.25% and 0.61% for $H \to \mu\tau$ and $H \to e\tau$ respectively [68]. In particular, the result of the $H \to e\tau$ search is now the leading upper limit on $H \to e\tau$ decays.

The results are also interpreted as a constraint on the modulus of the off-diagonal elements of the Yukawa couplings between Higgs boson and leptons. These limits are illustrated in figure 6.14. The limit are $\sqrt{|Y_{e,\tau}|^2 + |Y_{\tau,e}|^2} < 0.00197$ and $\sqrt{|Y_{\mu,\tau}|^2 + |Y_{\tau,\mu}|^2} < 0.00152$

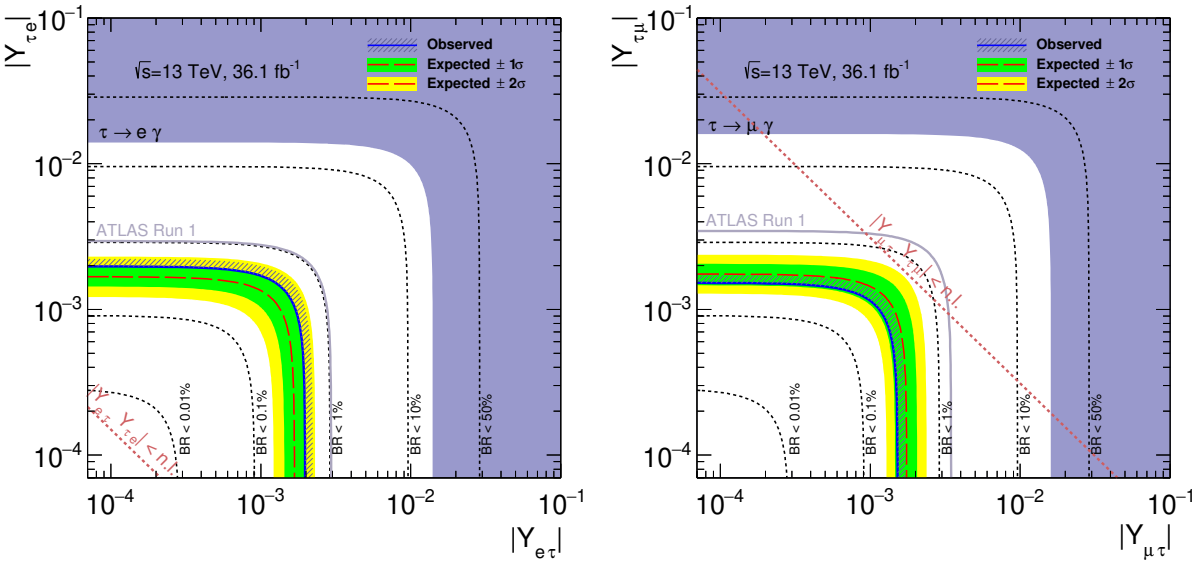


Figure 6.14: The upper limits on $|Y_{e,\tau}|$, $|Y_{\tau,e}|$ (left) and $|Y_{\mu,\tau}|$, $|Y_{\tau,\mu}|$ (right). Contours show constraints for different BR exclusions as well as the expected and observed limits. Also shown is a comparison with respect to the result from 7+8 TeV collision datasets [71, 72] as well as the most stringent limit from $\tau \to \ell\gamma$ searches [65]. The diagonal line represents the naturalness limit to retain preserve the mass hierarchy $|Y_{\ell,\tau}Y_{\tau,\ell}| < m_\tau m_\ell/\nu^2$.[65].

6.8 Conclusion and Future Prospects

A search for charged lepton flavour violation decays of $H \to \mu\tau$ and $H \to e\tau$ is presented and has found no such decays. The multivariate approach has been implemented via a BDT to more optimally extract the signal. The BDT has been shown to be effective in producing sensitivity which is competitive with respect to the CMS limits and this is driven by the BDT trained in the non-VBF category in the hadronic channel. The 95% confidence level upper limits placed on the branching ratio of $H \to e\tau$ and $H \to \mu\tau$ are 0.47% and 0.28% for the $H \to e\tau$ and $H \to \mu\tau$ decays respectively. The limit placed on the off-diagonal $Y_{e\tau}$ is currently the strongest limit for direct $H \to e\tau$ decays.

As the search is systematically limited, future studies on the LFV searches need to be directed in reducing the sources of systematic uncertainty. Additionally further work can be done to optimise the BDT training by attempting to train on a closer dataset to that of the signal region (i.e. train on the *tight* dataset). A re-optimisation of the input selection may also be useful.

7

Search for CP asymmetry with $H \to \tau\tau$ Decays

The motivations for studying CP violation in $H \to \tau\tau$ decays were presented in section 2.4.1. The measurement of the CP state in $H \to \tau\tau$ decays is best attempted with fully hadronic final states. As there is insufficient statistics for a precise measurement at the time of writing, a feasibility study is presented for the use of a multivariate classifier [147] rather than a traditional acoplanarity approach [49]. This approach aims to strengthen the sensitivity of the measurement by increasing the useable fraction of τ decays. This work has been published in [4].

7.1 Constructing a CP Sensitive Observable

Methods for constructing a CP sensitive observable is well established in the following [48–52, 148–150]. These methods involve reconstructing the decay plane of each τ lepton and measuring the angle between the two planes known as the acoplanarity angle (ϕ^*_{CP}).

In an extended non-CP-conserving Lagrangian, the Yukawa interaction between the Higgs boson and a pair of τ leptons can be expressed at tree-level as the sum of a CP-even and CP-odd terms

$$\mathcal{L}_{int} = g_\tau \overline{\tau}(\cos\phi_\tau + \sin\phi_\tau i\gamma_5)\tau h, \tag{7.1}$$

where ϕ_τ parameterises the mixing of couplings between CP even and CP odd Higgs fields.

Examination of the Higgs boson decay width reveals a dependence on to the mixing angle is evident with respect to the spin components of the τ leptons in the plane transverse to the τ momenta:

$$\Gamma(h_{mix} \to \tau^+\tau^-) \sim 1 - s_\parallel^{\tau^+} s_\parallel^{\tau^-} + s_\perp^{\tau^+} R(2\phi_\tau) s_\perp^{\tau^-} \tag{7.2}$$

where R is a rotation in the plane transverse to the direction of the τ momenta in the Higgs boson rest frame. This rotation matrix is a function of the mixing angle [50], the only such term. The $s_\parallel^{\tau^\pm}$, $s_\perp^{\tau^\pm}$ are the spin components of the τ in the directions transverse and parallel

the τ momenta in the Higgs boson rest frame, respectively. This dependence on the mixing angle is thus encoded in the transverse spin correlations between the two τ leptons.

These transverse spin correlations constrain the angular distributions of outgoing visible and invisible τ decay products. The particular bias in angular distributions are dictated by whether the coupling is CP-even or CP-odd. In the Higgs rest frame the differential decay width can be expressed as

$$d\Gamma(h_{mix} \to \tau^+\tau^-) \propto 1 - b(E_+)b(E_-)\frac{\pi^2}{16}(\cos(\phi^*_{CP} - 2\phi_\tau)) \tag{7.3}$$

where ϕ^*_{CP} is the acoplanarity angle, the angle between decay planes of the two τ, and b are spectral functions of each τ lepton in a given decay mode [151]. Measuring the differential rate with respect to ϕ^*_{CP}, produces a sinusoidal modulation whose phase encodes the mixing angle ϕ_τ. This definition, alone, however is not sufficient to generate a modulation which differs between different mixing parameters. Integration over the full phase space will leave a flat ϕ^*_{CP} distribution if events are not properly categorised. To demonstrate this, consider the matrix element generating the angular distribution:

$$w_{\text{spin}} \propto 1 + \sum_{i,j} R_{i,j} h^i h^j \tag{7.4}$$

where h_i and h_j are polarimetric vectors [152] and $R_{i,j}$ is defined as

$$R = \text{diag}(-1, +1, +1) \text{ and } \text{diag}(-1, -1, -1) \tag{7.5}$$

for scalar and pseudoscalar Higgs bosons respectively. The polarimetric vectors represent the most likely τ spin direction given a configuration of the decay products. For decays to ρ mesons, which produce one charged and neutral pion - one of the most sensitive channels, this vector is expressed as:

$$h^i = f(2(\mathbf{q} \cdot \mathbf{p}_\nu)\mathbf{q} - |\mathbf{q}|^2\mathbf{p}_\nu), \tag{7.6}$$

where f is a normalisation function dependent on $\mathbf{p}_\nu$ and $\mathbf{q}$ which is the difference between pion four vectors $\mathbf{p}_{\pi^\pm} - \mathbf{p}_{\pi^0}$. The equation 7.6 indicates that events must be separated in order to avoid cancellation in modulations when accounting for the spin correlations.

An asymmetry variable y is defined as

$$y_\pm = \frac{E_{\pi^\pm} - E_{\pi^0}}{E_{\pi^\pm} + E_{\pi^0}} \tag{7.7}$$

for each τ lepton, where E is the lab frame energy of the pions in the decay. Events with $y_+y_- > 0$ will have a phase shifted 90° with respect to events with $y_+y_- < 0$.

Categorised events according to the sign of product y_+y_- will thus prevent cancellation of the ϕ^*_{CP} modulation when integrating over all events, leaving sensitivity to the mixing angle.

7.1.1 Decay Mode Challenges

The determination of the acoplanarity angle has unique challenges and benefits depending on the decay mode of the τ lepton. The sensitivity to the CP state of the Higgs boson is influenced by the spin analysing power of the τ decay, which differs greatly between the decay modes. Of the two broad categories of τ decays (to either leptonic or hadronic decay products) the hadronic mode has the stronger spin analysing power.

The sensitivity to the transverse spin correlations for individual decay modes is dependent on a mixture of theoretical and experimental considerations. In the various decay modes of the τ lepton, the dependence of the transverse spin of a τ manifests in a number of angles. For example the angle between the τ vector and the $\pi^\pm$ for direct decay to $\pi^\pm\nu$. For more complex decays, such as a decay to $\rho^\pm$ and $a_1^\pm$, the corresponding cascade has the effect of diluting the sensitivity of a larger number of angles [153] The more complex the more angles one must analyse.

From an experimental perspective, the presence of a neutrino in the final state creates an issue in being able to define the true τ momenta direction. This creates issues using the single charged particle modes, such as $\tau^\pm \to \pi^\pm\nu$, without the use of additional information [52, 148–150].

Thus the focus of this chapter will be to explore the possibility of measurement with decays via $\tau^\pm \to \rho^\pm\nu$ and $\tau^\pm \to a_1^\pm\nu$. The following sections will review the different methods available for reconstructing acoplanarity angles based on these decay modes. Issues relating to experimental effects are described together with an overview of attempts to alleviate these problems.

$\rho^\pm$ Decay Mode

The simplest, most robust decay mode to study is the case where the τ decays via a $\rho^\pm$ meson. The full chain is $\tau^\pm \to \rho^\pm\nu \to \pi^\pm\pi^0\nu$. This decay mode accounts for approximately 25% of all τ lepton decays, constituting the single largest branching fraction for a single decay mode.

In this case, the acoplanarity angle has been shown to be sensitive to the mixing angle via some simple kinematic considerations [49, 50]. For this decay mode a plane can easily be constructed from the momenta of the charged and neutral pions for each of the decay τ leptons. This angle would ideally be measured in the Higgs boson rest frame, however due to the presence of neutrinos in the decay, the frame can not be fully reconstructed with a good degree of precision. Instead, the rest frame of the $\rho^\pm - \rho^\mp$ system is used: this is a

close approximation to the Higgs boson rest frame. The angle between these two planes is defined as the acoplanarity angle. A schematic is presented in 7.1.

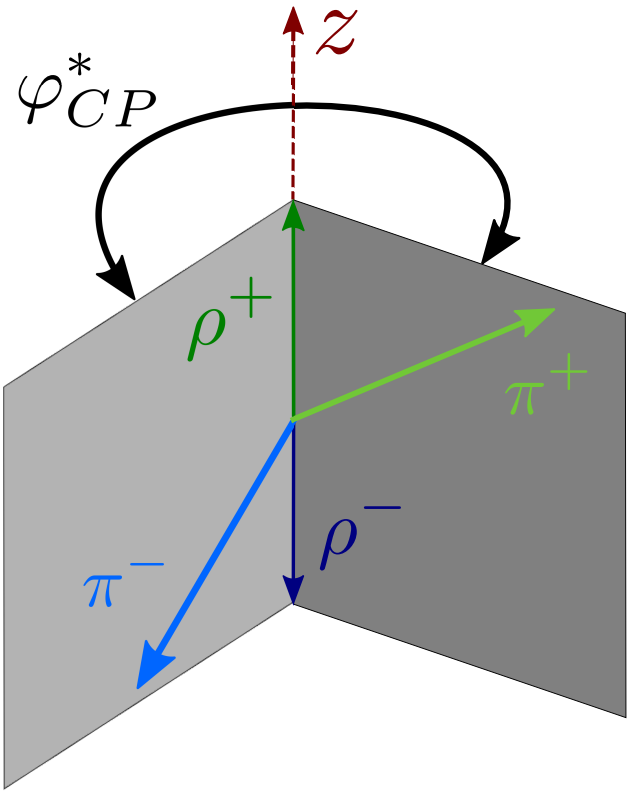
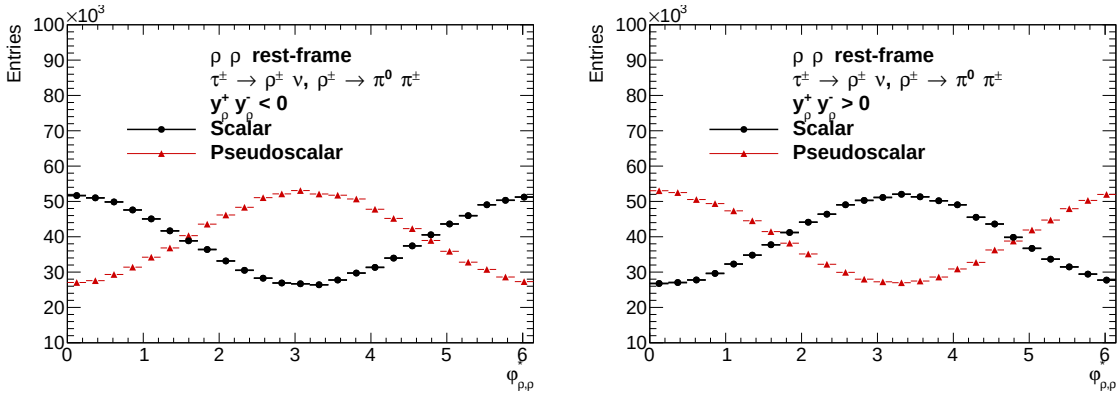


Figure 7.1: A schematic of the calculation of the ϕ^*_{CP} angle for in the case where the two τ decay via the chain $\tau^\pm \to \rho^\pm\nu \to \pi^+\pi^0\nu$. This schematic depicts the pions from the τ decays in the zero momentum frame of the $\rho^\pm - \rho^\mp$ system. The system here is aligned such that the $\rho^\pm$ are aligned in the z axis.

The acoplanarity angle is expected to produce a sinusoidal modulation which is shifted in phase with respect to differing mixing angles (see figure 7.2). Owing to the simplicity of the decay, this angle is defined in an unambiguous and fairly robust manner.

Experimentally speaking, this channel is rather free from large reconstruction effects. Aside from the lack of accurate neutrino reconstruction, which impacts all τ decay modes, this mode suffers most from a large cross-contamination from decays with multiple neutral pions. This is caused by the relatively poor reconstruction of neutral pions (see section A.1 for more details).



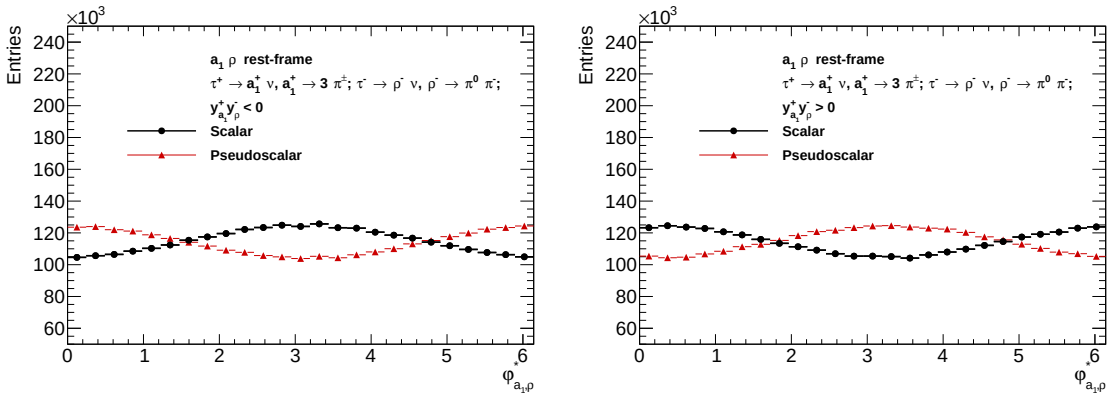
(a) $y_+ \cdot y_- < 0$ (b) $y_+ \cdot y_- > 0$

Figure 7.2: Distributions of acoplanarity angles of $H \to \tau\tau$ decays in the $\rho - \rho$ decay mode [147].

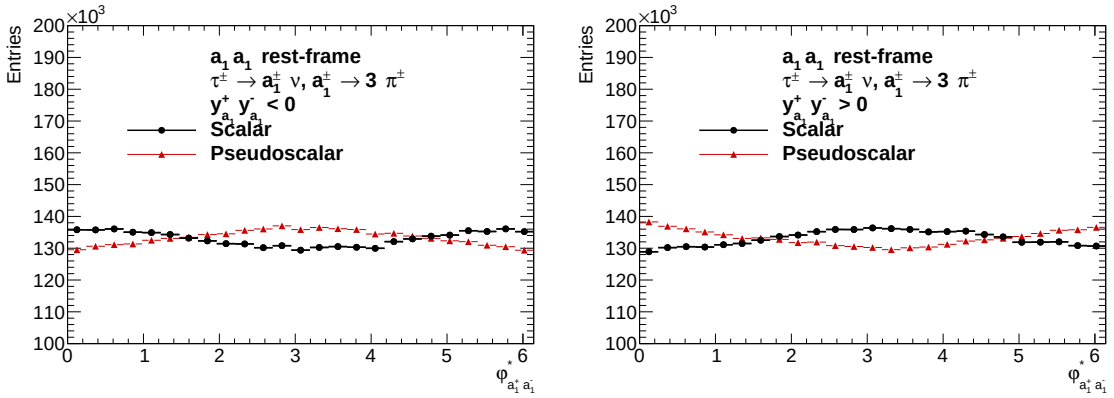
Three Pion Decays

The approach of simply taking a single acoplanarity angle is not effective for more complex decays via cascade decays. One such cascade decay proceeds through the chain $\tau^\pm \to a_1^\pm \nu \to \pi^\pm \rho^0 \nu \to \pi^\pm \pi^\pm \pi^\mp \nu$. Also possible is the case where ρ^0 decays to a pair of *neutral* pions however this will not be discussed further as neutral pions are more challenging to to reconstruct than charged pions.

In the complex decay to three charged pions, it is not as evident how exactly to define the acoplanarity angle such that all the CP sensitive information is captured in a single observable. The sensitivity of the transverse spin is diluted over several more angles as the decay chain becomes more complex [153]. Figure 7.3 demonstrates this effect. The more complex decays which involve a_1 resonances have a much smaller amplitude of the modulation which scalar and pseudoscalar hypotheses compared to figure 7.2 . This results in the observable performing much worse when resolving the mixing angle ϕ_τ for more complex decays.



(a) $\tau^{\pm} \to \rho^{\pm}\nu_{\tau}, \tau^{\mp} \to a_1^{\mp}\nu_{\tau}$



(b) $\tau^{\pm} \to \rho^{\pm}\nu_{\tau}, \tau^{\mp} \to a_1^{\mp}\nu_{\tau}$

Figure 7.3: Distributions of acoplanarity angles of $H \to \tau\tau$ decays [147]. Plots right (left) show events where $y_+ \cdot y_-$ is less than (greater than) zero.

In spite of the complexity of the $a_1 - a_1$ decay, the inclusion of these decays would yield a significant increase in the useable fraction of $H \to \tau\tau$ decays (from 6.5% with only decays via the ρ to 11.9% with the inclusion of a_1 decays). For $a_1 - a_1$ decays though, it is possible to form 16 acoplanarity angles with 8 asymmetry y variables. The multi-dimensionality of the problem suggests the need for a more comprehensive approach. Here, neural networks are employed to build a new CP observable. The goal is build classifiers which will be utilised for any τ decay mode and maximise the sensitivity to the CP mixing angle by exploiting simple and robust inputs.

7.2 Neural Network Approach

Neural networks are a form of supervised learning algorithms similar to BDTs discussed in section 6.4.1. While BDTs are essentially decision tree structures, neural networks take a more complex form. Inspired by connections in the human brain, NN are structures of connected nodes. Activation of pathways between nodes by a stimulus induces a certain response. In a similar sense, neural networks take input features and connects them to one or more layer(s) of nodes which are connected via a combinations of linear functions of weights and biases.

The inter-connectedness of such a network allows very complex observables to be constructed which is not easily realisable with a BDT structure. The complexity and high dimensionality make the neural network approach ideal for the measurement attempted here.

Training the neural network involves iterating over a given sample and continually modifying the weights in order to minimise an objective loss function. Recalling the description of BDTs in section 6.4.1, a smooth differentiable gradient function is then constructed from the weights and biases. The backpropagation process, whereby weights and biases are updated according to the gradient, continually improves the network for each iteration.

The exact sample, structure of the NN and loss function will be detailed in the following sub-section.

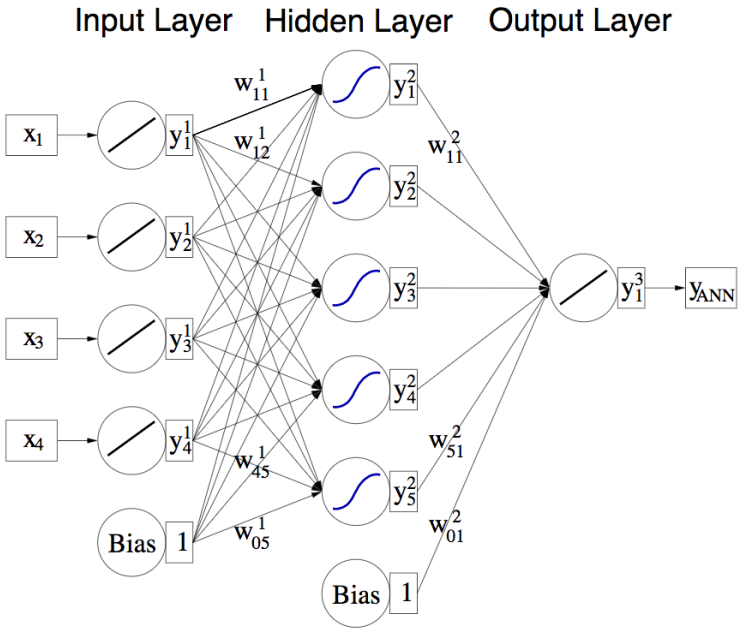


Figure 7.4: A schematic of a single layer NN, each connection contains a weight and bias which is optimised in the training process. The single output is the classifier score.

7.2.1 Inputs and NN Setup

To implement this approach, MC samples were generated in order to train and to test the effectiveness of the NN approach. The simulations modelled $H \to \tau\tau$ decays which were produced via ggH in $\sqrt{s} = 13$ TeV proton-proton collisions. Between 2 and 5 million $H \to \tau\tau$ events were generated for each decay mode using `Pythia 8.2`. The decays were simulated using the `TAUOLA` library [100] and the `TauSpinner` [154] package was used to calculate the event spin weights for the scalar and pseudoscalar hypotheses. Of these events, after separating events for training and testing, as well as after the application of ATLAS detector acceptance selections, approximately 500000 events were utilised for training the neural networks.

Combinations of the following classes of inputs, all expected to provide some degree of discrimination power, were used in training the neural networks:

- ϕ^* - The acoplanarity angle, fundamental CP sensitive variable which is defined as the angle between two reconstructed decay planes [49]. All possible combinations between pairs and triplets of pions are accounted for.

- y - The separating variable used to categorise events to avoid cancellation of modulations in ϕ^* [49]. The definition for decays to $\rho^{\pm}$ resonances was detailed in equation 7.7. A modification is required for decays of $a_1 \to \rho^0\pi^{\pm}$ due to the large mass of the ρ^0 resonance. For these decays, y is defined as $\frac{E_{\rho^0}-E_{\pi^{\pm}}}{E_{\rho^0}+E_{\pi^{\pm}}} - \frac{m^2_{a_1}-m^2_{\pi^{\pm}}+m^2_{\rho^0}}{2m^2_{a_1}}$ [147].

- m_i- Invariant masses for pairs or triplets of pions. This is especially useful for τ decays proceeding via an a_1 meson cascade decay as it allows for determination of the likelihood at which a pion pair reproduces the ρ^0 meson.

- 4-vectors - The four-momenta of the outgoing pions. It was shown in [147] that these variables, can be effectively used in place of the other inputs. These inputs were boosted into the rest frame of the visible decay products and then rotated such that one of the reconstructed taus aligns along the positive z-axis (see figure 7.5 for a schematic). This procedure was implemented avoid the neural network re-learning trivial symmetries such as rotations and boosts. In principle, this class of inputs contains all that is required to reconstruct the other variables mentioned.

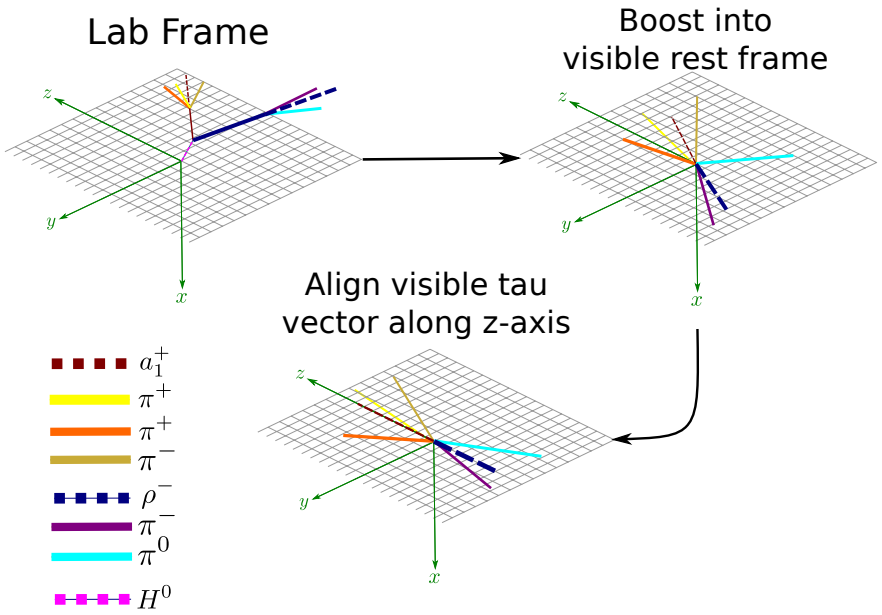


Figure 7.5: A schematic of the procedure of boosts and rotations needed to make the system of 4-vector inputs uniformly aligned to reduce trivial symmetries.

With the inputs calculated, neural networks were trained using different combinations of input variables within the KERAS [155] package interfaced with the TENSORFLOW machine learning package [156]. The neural networks consisted of six dense (fully connected) layers of 300 nodes with a ReLU activation function for each non-terminal node. The activation function dictates when an combination of inputs has sufficiently large weight as to "activate" the node for further use (see [157] for details on the ReLU activation and it's use in neural networks). The final layer is then connected to a single terminal node (the classifier score) which utilises a sigmoid to compress the score to the range $[0, 1]$ so that the value can be interpreted as a probability. The Adam algorithm [158] is utilised as the optimisation algorithm (which is based off of a stochastic gradient descent method) and the NN was trained over 50 iterations. The networks were trained to provide a probability of an event originating from a scalar or pseudoscalar Higgs decay. A minimal amount of dropout [159] was implemented to prevent overtraining whilst still retaining sensitivity. The AUC was taken as the metric for the classification power of each network.

7.2.2 Limitations of the Neural Network Approach

The simulation used to estimate the sensitivity of the NN approach is a particle level description of the $H \to \tau\tau$ final state. The detection and reconstruction of the final state with the ATLAS detector is not simulated. To get an realistic estimate of the sensitivity of this method, one must check whether it is impacted by specific experimental as well as theoretical effects. The detector resolution limits the precision in the reconstruction of the four-momenta of the outgoing pions, and this may reduce the sensitivity of the NN approach. To address this, simple Gaussian smearings (based on detector resolutions representative of the ATLAS detector [160, 161]) were applied the simulated particle momenta to simulate to assess the potential impact. Additionally, the impact on the sensitivity of the NN method from systematic uncertainties related to the theoretical modelling of τ decays is also presented.

7.2.3 Detector Resolution Effects

Training against particle level simulation provides a baseline for the sensitivity of the approach, however this does not encapsulate the limitations due to the detector resolution. The degradation due to limited detector resolution is assessed by smearing the original ("ideal") MC to produce "smeared" samples which are more representative of the events reconstructed with the ATLAS detector.

Table 7.1 presents the AUC score calculated by applying the NN trained on their respective MC (either generator level or smeared) to their respective test dataset for combinations of input features. The quoted statistical and smearing uncertainties are as follows:

- Train a NN on either generator level or smeared MC
- Duplicate the test dataset multiple times (of order 1000 times) to create an ensemble of dataset.
- For each dataset in the ensemble perform a variation:
  - For the statistical uncertainty, perform a bootstrap sampling (sampling with replacement from the dataset) [162]. This is done for either ideal or smeared MC
  - For the smearing uncertainty, vary the random seed used to smear the sample. This is only performed on ideal MC to produce variations of the smearing.
- Reapply the trained NN on each of the datasets in the ensemble and calculate the AUC score
- Take the mean and width of the distribution of the AUC scores as the central value and the uncertainty respectively.

Features				Ideal ± (stat)	Smeared ± (stat) ± (syst)	From [147]
ϕ^*	4-vec	y_i	m_i			
$a_1 - \rho$ Decays						
✓	✓	✓	✓	0.6035 ± 0.0005	$0.5923 \pm 0.0005 \pm 0.0002$	0.596
✓	✓	✓	-	0.5965 ± 0.0005	$0.5889 \pm 0.0005 \pm 0.0002$	-
✓	✓	-	✓	0.6037 ± 0.0005	$0.5933 \pm 0.0005 \pm 0.0003$	-
-	✓	-	-	0.5971 ± 0.0005	$0.5892 \pm 0.0005 \pm 0.0002$	0.590
✓	✓	-	-	0.5971 ± 0.0005	$0.5893 \pm 0.0005 \pm 0.0002$	0.594
✓	-	✓	✓	0.5927 ± 0.0005	$0.5847 \pm 0.0005 \pm 0.0002$	0.578
✓	-	✓	-	0.5819 ± 0.0005	$0.5746 \pm 0.0005 \pm 0.0002$	0.569
$a_1 - a_1$ Decays						
✓	✓	✓	✓	0.5669 ± 0.0004	$0.5657 \pm 0.0004 \pm 0.0001$	0.573
✓	✓	✓	-	0.5596 ± 0.0004	$0.5599 \pm 0.0004 \pm 0.0001$	-
✓	✓	-	✓	0.5677 ± 0.0004	$0.5661 \pm 0.0004 \pm 0.0001$	-
-	✓	-	-	0.5654 ± 0.0004	$0.5641 \pm 0.0004 \pm 0.0001$	0.553
✓	✓	-	-	0.5623 ± 0.0004	$0.5615 \pm 0.0004 \pm 0.0001$	0.573
✓	-	✓	✓	0.5469 ± 0.0004	$0.5466 \pm 0.0004 \pm 0.0001$	0.548
✓	-	✓	-	0.5369 ± 0.0004	$0.5374 \pm 0.0004 \pm 0.0001$	0.536

Table 7.1: AUC scores for NNs trained on various combinations of input features (those used are marked with a ✓). The columns represent the results for training (and application) on "Ideal" and "Smeared" samples. The last column presents AUC scores with [147] where no detector resolution effects were implemented for similarly trained networks where the network structure was static across all trained networks (i.e. no optimisation).

The results from table 7.1 demonstrate only a very small loss in sensitivity ($\sim 1\%$) by training on the smeared samples. This trend is consistent across both of the decay modes and the different combinations of input features indicating the NN approach is robust against the detector resolution in reconstructing the pions from the τ decays.

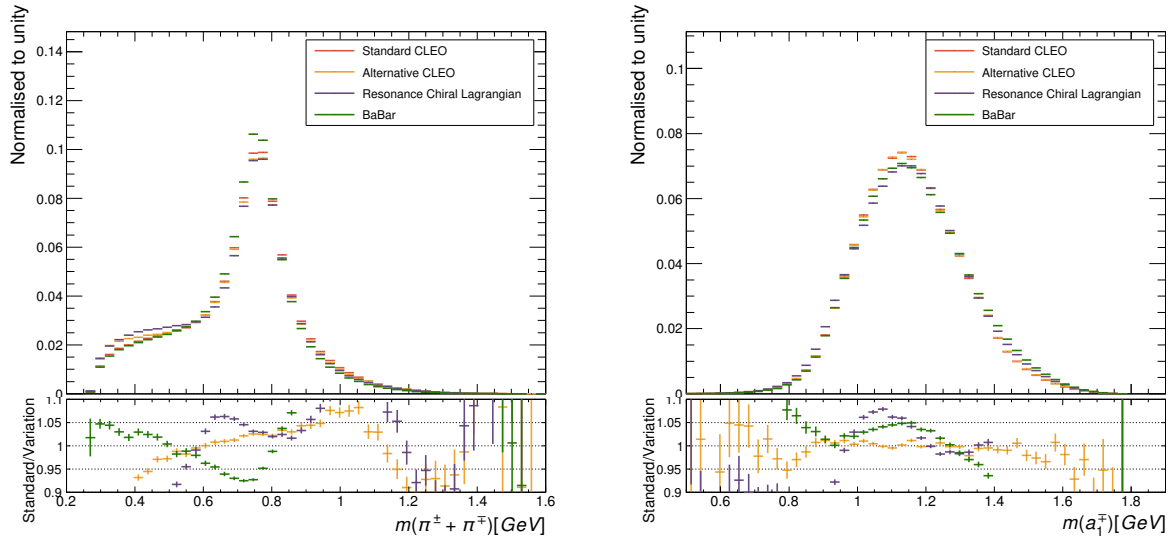
7.2.4 Systematic Uncertainty due to τ Modelling

One source of systematic uncertainty which could impact the sensitivity of the NN, is the modelling of the τ decays. The `TAUOLA` library [100] models the τ decays using data-driven parameterisations using information from low-energy collider experiments such as CLEO [163] and BaBar [164]. Decays via the a_1 resonance may be sensitive to the modelling of spin effects that are propagated to the vector resonances which make it difficult to measure. Thus, through the spin correlations between τ leptons (which are sensitive to the CP state of the Higgs boson), the modelling of the decays via the a_1 may be a crucial systematic which impacts any CP measurement, not only the NN approach.

To test this potential effect, various parameterisations of the τ modelling were tested against NNs trained on the default CLEO based hadronic current parametrisation, the default for the `TAUOLA` library. The following variations on the hadronic current modelling are used for testing:

- Standard CLEO (STD) - the default parameterisation [163] based on the Kuhn-Santamaria (KS) model [99]. This library is used to train the NN which is applied to the other samples.
- Alternative CLEO (ALT) - a variation of the STD which described in [165] which is taken from an isospin rotation from $\pi^0\pi^0\pi^-$ to the $\pi^-\pi^-\pi^+$ channel.
- BaBar (BBR) - the same KS model as with the STD current but using measurements from the BaBar collaboration [166].
- Resonance Chiral Lagrangian (RχL) - a fundamentally different model to the KS detailed in [167].

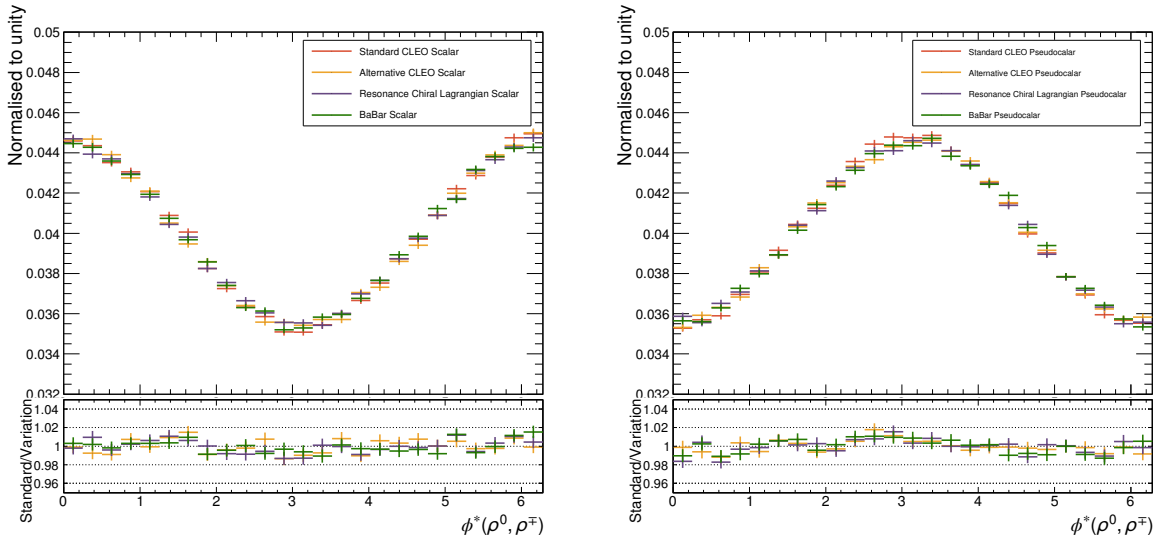
Figure 7.6 demonstrates the effect of the different parameterisations on the mass (which typically is used to determine good modelling of the τ decays). Figures 7.7 and 7.8 shows the corresponding effect on the acoplanarity angle. Despite the acoplanarity showing little significant variation in the acoplanarity angle it is important that the correlations be checked in the NN. In order to check if the variations cause a loss in sensitivity, NNs which were trained on MC generated with the standard CLEO parameterisation are applied to MC generated with the other three variations.



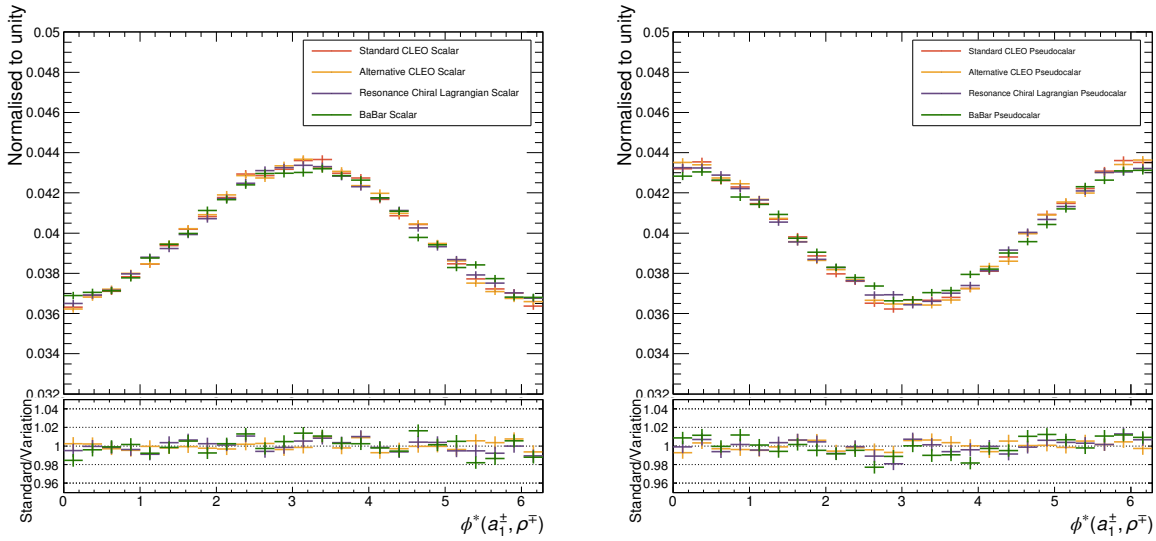
(a) Mass of $\pi^{\pm}\pi^{\mp}$

(b) Mass of three pions

Figure 7.6: Invariant masses constructed from $\tau^{\mp} \to a_1^{\mp}\nu \to 3\pi^{\mp}\nu$ decays using different parameterisations of the hadronic current. Ratios between the alternative (RχL, ALT, BBR) and the baseline (STD) parameterisation are given in the lower panels.

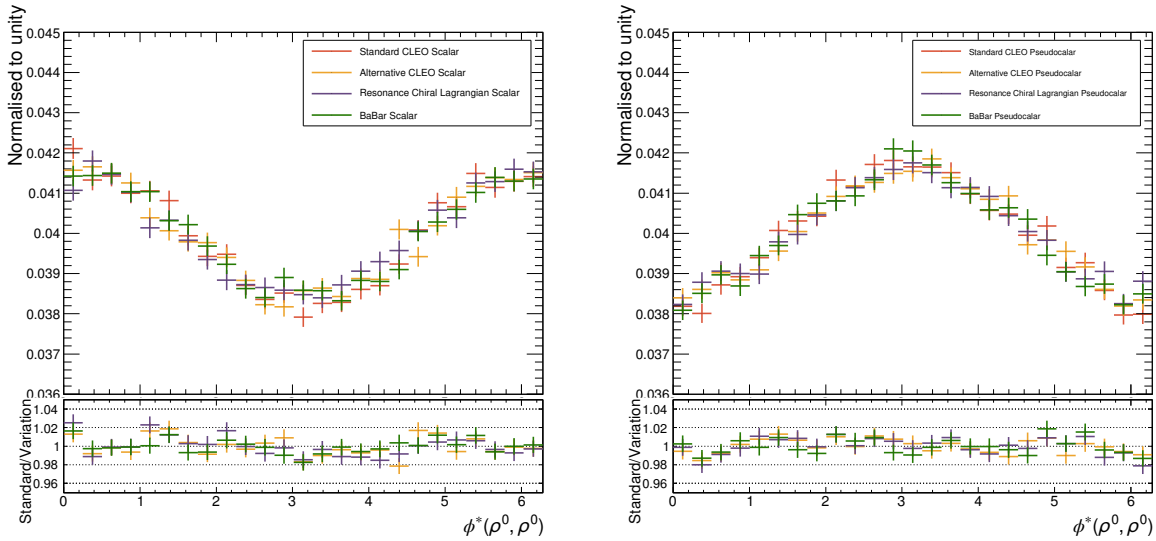


(a) Plane for $\tau^\mp$ from $\pi^\mp - \pi^0$, plane for $\tau^\pm$ from $\pi^\pm - \pi^\mp$

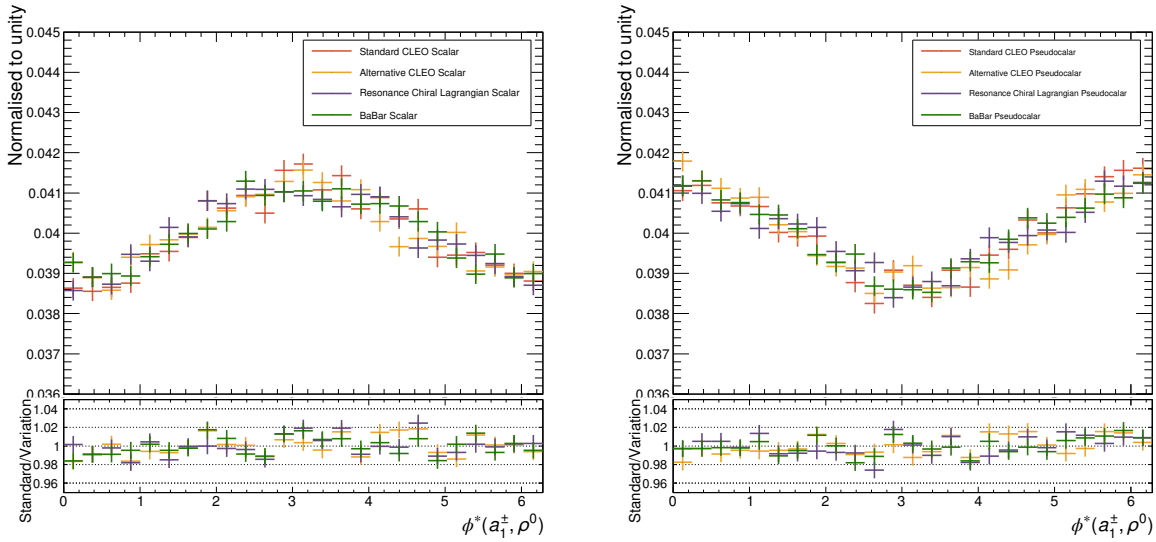


(b) Plane for $\tau^\mp$ from $\pi^\mp - \pi^0$, plane for $\tau^\pm$ from $\pi^\pm - \rho^0$

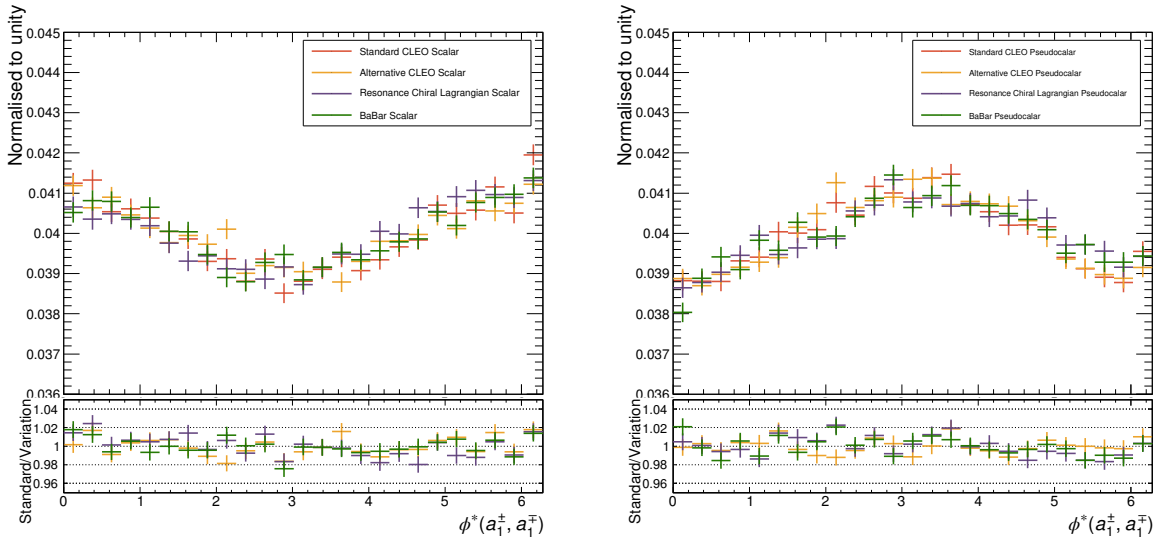
Figure 7.7: A comparison of decay rates as a function of the acoplanarity angles calculated for $H \to \tau\tau \to \rho^\mp\nu\, a_1^\pm\nu$ using different parameterisations (STD, RχL, ALT, BBR) for events with $y_+ \cdot y_- > 0$. Ratios between the alternative (RχL, ALT, BBR) and the baseline (STD) parameterisation are given in the lower panels. Note that ρ^0 indicates the vector used to construct the plane was formed from a $\pi^\pm - \pi^\mp$ pair. Each row contains the distributions of the acoplanarity angle for scalar (left) and pseudoscalar (right) hypotheses.



(a) Plane for $\tau^{\mp}$ from $\pi^{\mp} - \pi^{\pm}$, plane for $\tau^{\pm}$ from $\pi^{\pm} - \pi^{\mp}$



(b) Plane for $\tau^{\mp}$ from $\pi^{\mp} - \pi^{\pm}$, plane for $\tau^{\pm}$ from $\pi^{\pm} - \rho^0$



(c) Plane for $\tau^{\mp}$ from $\pi^{\mp} - \rho^0$, plane for $\tau^{\pm}$ from $\pi^{\pm} - \rho^0$

Figure 7.8: A comparison of decay rates as a function of the acoplanarity angles calculated for $H \to \tau\tau \to a_1^{\mp}\nu\, a_1^{\pm}\nu$ using different parameterisations (STD, RχL, ALT, BBR) for events with $y_+ \cdot y_- > 0$. Ratios between the alternative (RχL, ALT, BBR) and the baseline (STD) parameterisation are given in the lower panels. Each row contains the distributions of the acoplanarity angle for scalar (left) and pseudoscalar (right) hypotheses.

The results in table 7.2 show that the fluctuations in the AUC score (and hence the sensitivity) are within two or three times the quoted statistical uncertainty from table 7.1. Ultimately, the loss in sensitivity due to training on events with a more realistic detector simulation, impacts the deep learning approach to a greater degree than the modelling of the τ decay.

Features				STD	RχL	ALT	BBR
ϕ^*	4-vec	y_i	m_i				
$a_1 - \rho$ Decays							
✓	✓	✓	✓	0.604	0.604	0.603	0.603
✓	✓	✓	-	0.597	0.596	0.596	0.597
✓	✓	-	✓	0.604	0.604	0.604	0.604
-	✓	-	-	0.597	0.596	0.596	0.595
✓	✓	-	-	0.597	0.596	0.596	0.595
✓	-	✓	✓	0.593	0.593	0.593	0.593
✓	-	✓	-	0.582	0.579	0.580	0.578
$a_1 - a_1$ Decays							
✓	✓	✓	✓	0.567	0.563	0.564	0.564
✓	✓	✓	-	0.560	0.555	0.557	0.556
✓	✓	-	✓	0.568	0.564	0.566	0.566
-	✓	-	-	0.562	0.557	0.559	0.559
✓	✓	-	-	0.562	0.557	0.559	0.559
✓	-	✓	✓	0.547	0.546	0.547	0.545
✓	-	✓	-	0.537	0.534	0.535	0.533

Table 7.2: AUC score for NNs trained with $a_1 - \rho$ and $a_1 - a_1$ decays of ditau system using events modelled with the STD and then tested on events generated with alternative parameterisations. The test was performed on events simulated at generator level.

7.3 Conclusion and Outlook

A novel method has been developed to extend the potential measurement of a Higgs boson CP mixing angle to use three pion τ decays. This neural network approach can be applied to decay modes involving resonant decays to $\rho^{\pm}$ and $a_1^{\pm}$ mesons. This chapter has shown that the inclusion of smearing and theoretical systematics in the τ decay model has shown to be of little impact to this method. It is hoped that this methodology can be adapted to be a wholly encompassing method to consistently measure the Higgs boson CP across all τ decay modes.

The first steps in producing a measurement in ATLAS is to recheck these studies with a full detector simulation, properly accounting for more subtle reconstruction effects which can not be readily accessed in studies external to the collaboration. A training across multiple decay modes, carefully combining and accounting for all possibilities, would then be the next step. One key challenge is to properly reconstruct the impact parameters and potentially secondary vertices as inputs for more nuanced decay modes such as the single pion.

Further development to the reconstruction of the neutrino four-momentum would provide the most substantial improvement in sensitivity as full sensitivity to the mixing angle could be achieved. Unfortunately, the performance of the MMC algorithm is currently inadequate. Development of a new method of neutrino reconstruction would be the key to recovering missing components of the sensitivity to the CP state of the Higgs boson in ditau decays.

8

Conclusion

The discovery of the Higgs boson in 2012 has opened particle physics to tests of new physics using Higgs boson decays. Measurements of expected SM couplings validate the Higgs mechanism in its role in the SM whereas searches for new couplings which violate SM symmetries are probes of new physics. Using 36.1 fb^{-1} of $\sqrt{s} = 13$ TeV data, a measurement of SM couplings to τ leptons and a search for lepton flavour violating $H \to e\tau$ and $H \to \mu\tau$ decays are both performed. This thesis outlined both measurement and search as well as a feasibility study for measuring the CP state of the Higgs boson.

Prior to any measurement or search, a calibration of the τ energy scale was performed with 3 fb^{-1} of $\sqrt{s} = 13$ TeV data. Using a tag-and-probe technique looking at $Z \to \tau_\mu \tau_{had}$ decays, an insitu calibration was performed. A simple OS-SS background estimate was employed with simple selections to define a pure sample of $Z \to \tau_\mu \tau_{had}$ decays. Following this, a binned χ^2 test was performed to assess the need for a correction to the τ energy scale. It was found that the τ energy scale was well calibrated with a required shift being determined to be $\alpha = -0.7 \pm 0.8(\text{stat}) \pm 1.2(\text{syst})$ and $\alpha = -3.6 \pm 1.2(\text{stat}) \pm 3.0(\text{syst})$ for 1p and 3p, respectively. This correction is consistent with a shift parameter of 0%.

With a well calibrated τ algorithm, a measurement of the elusive $H \to \tau\tau$ process was performed. A completely reoptimised analysis strategy for semileptonic $H \to \tau\tau$ decays was implemented to measure this important coupling. This analysis strategy relied on selections with kinematic properties of the Higgs boson signal rather than a multivariate technique. This strategy split existing categories into signal regions with varying signal sensitivities in order to retain maximum signal acceptance whilst improving the overall sensitivity of the measurement. With a well established background estimate, the semileptonic channel proved to be a crucial key in this measurement. The combination of all channels involved the integration of over 200 parameters with 13 signal regions and 6 control regions. Extensive testing ensured that a cohesive and robust statistical model was built. Ultimately this measurement proved that the $H \to \tau\tau$ coupling existed and was consistent with a SM expectation. The measured cross-section was $\sigma_{H\to\tau\tau} = 3.71 \pm 0.59(\text{stat})^{+0.87}_{0.74}(\text{syst})$ pb with respective cross-sections of $\sigma^{ggH}_{H\to\tau\tau} = 3.0 \pm 0.13$ and $\sigma^{VBF}_{H\to\tau\tau} = 0.237 \pm 0.006$ pb for the individual ggH and VBF production modes. The discovery of this mode in ATLAS along with

all measurements being consistent with the SM conclusively proves the Higgs boson's role in mass generation for τ leptons. It also stands as a key indicator of the Higgs boson's role in fermion mass generation as a whole.

Building from the measurement of $H \to \tau\tau$ decays, a search for lepton flavour violating $H \to e\tau$ and $H \to \mu\tau$ decays was presented. Intriguing excesses in the $\sqrt{s} = 8$ TeV dataset motivated a search for these decays which could herald new physics in the Higgs boson sector. For this search an identical dataset, background estimation technique, systematic uncertainty calculation and similar particle candidate selections and statistical model were reused from the $H \to \tau\tau$ measurement. The main driver of sensitivity was the development of multivariate BDT algorithms in order to extract maximum sensitivity from a process expected to have less than 1% of the Higgs boson branching fraction. Rather than relying on simple event selections in the case of the $H \to \tau\tau$ decays, a BDT discriminator would provide the background rejection, in particular for the non-VBF category of the τ_{had} channel. Differences in correlations of sensitive kinematic variables between background and signal processes were exploited in order to be maximally effective in discriminating the tiny signal from the overwhelming background. The non-VBF category of the hadronic τ decays, containing the bulk of the background in these searches is now driving the sensitivity of this search. No excesses were observed but upper limits on the branching fraction of $H \to e\tau$ and $H \to \mu\tau$ were set at 0.47% and 0.28%, respectively. The $H \to e\tau$ upper limit now leads the global limits for $H \to e\tau$ couplings.

The final test presented was a feasibility study for the measurement of the CP state of the Higgs boson. The $H \to \tau\tau$ decays provide a direct coupling to the Higgs boson which allow for mixing between scalar and pseudoscalar components at tree level without the need for higher order operators. Construction of an observable sensitive to this CP mixing effect is well established for simple mode such as the decay via a $\rho^{\pm}$ resonance, but this restricts the branching fraction to a measly 6%. By incorporating more complicated modes such as via the $a_1^{\pm}$ resonance, one is able to incorporate more of the $H \to \tau\tau$ branching fraction. A use of deep learning neural networks has been pioneered in order to cope with the high multiplicity final state of the $a_1^{\pm}$ resonance. Smeared MC simulation was used to train and test the effectiveness of the networks against detector resolution. It has been shown that the method to is indeed robust against these effects. In addition effects due to the model used to decay τ lepton was tested. Testing against MC simulation utilising alternative parameterisations of the τ_{had} decays demonstrated no sensitivity of the neural networks to these effects. It is hoped these techniques can be expanded upon to incorporate many more decay modes. Once sufficient data is collected, these techniques can be used to in order to conclusively prove the scalar nature of the Higgs boson.

Bibliography

[1] The ATLAS collaboration. "Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV". In: (2017).

[2] *Cross-section measurements of the Higgs boson decaying to a pair of tau leptons in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. ATLAS-CONF-2018-021. Geneva: CERN, June 2018. URL: http://cds.cern.ch/record/2621794.

[3] Morad Aaboud et al. "Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Submitted to: Phys. Rev.* (2018). arXiv: 1811.08856 [hep-ex].

[4] Elisabetta Barberio et al. "Deep learning approach to the Higgs boson CP measurement in $H \to \tau\tau$ decay and associated systematics". In: *Phys. Rev.* D96.7 (2017), p. 073002. DOI: 10.1103/PhysRevD.96.073002. arXiv: 1706.07983 [hep-ph].

[5] S. L. Glashow. "Partial Symmetries of Weak Interactions". In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[6] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[7] Abdus Salam and John Clive Ward. "Weak and electromagnetic interactions". In: *Nuovo Cim.* 11 (1959), pp. 568–577. DOI: 10.1007/BF02726525.

[8] Lyndon Evans and Philip Bryant. "LHC Machine". In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[9] G. Aad et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[10] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[11] Peter W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Phys. Lett.* 12 (1964), pp. 132–133. DOI: 10.1016/0031-9163(64)91136-9.

[12] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[13] Peter W. Higgs. "Spontaneous Symmetry Breakdown without Massless Bosons". In: *Phys. Rev.* 145 (1966), pp. 1156–1163. DOI: 10.1103/PhysRev.145.1156.

[14] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys. Rev. Lett.* 13 (1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585.

[15] T. W. B. Kibble. "Symmetry breaking in nonAbelian gauge theories". In: *Phys. Rev.* 155 (1967), pp. 1554–1561. DOI: 10.1103/PhysRev.155.1554.

[16] John Ellis. "Higgs Physics". In: arXiv:1312.5672. KCL-PH-TH-2013-49. LCTS-2013-36. CERN-PH-TH-2013-315 (Dec. 2013). 52 pages, 45 figures, Lectures presented at the ESHEP 2013 School of High-Energy Physics, to appear as part of the proceedings in a CERN Yellow Report, 117–168. 52 p. URL: http://cds.cern.ch/record/1638469.

[17] J. Goldstone. "Field Theories with Superconductor Solutions". In: *Nuovo Cim.* 19 (1961), pp. 154–164. DOI: 10.1007/BF02812722.

[18] *LHC Higgs Cross Section Working Group (2012-2013)*. accessed 2019-02-12. URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections.

[19] Georges Aad et al. "Search for the Standard Model Higgs boson produced in association with a vector boson and decaying into a tau pair in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Phys. Rev.* D93.9 (2016), p. 092005. DOI: 10.1103/PhysRevD.93.092005. arXiv: 1511.08352 [hep-ex].

[20] M. Aaboud et al. "Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector". In: *Phys. Lett.* B784 (2018), pp. 173–191. DOI: 10.1016/j.physletb.2018.07.035. arXiv: 1806.00425 [hep-ex].

[21] S. Dittmaier et al. "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables". In: (2011). DOI: 10.5170/CERN-2011-002. arXiv: 1101.0593 [hep-ph].

[22] Morad Aaboud et al. "Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector". In: *Phys. Lett.* B784 (2018), pp. 345–366. DOI: 10.1016/j.physletb.2018.07.050. arXiv: 1806.00242 [hep-ex].

[23] M. Aaboud et al. "Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment". In: *Phys. Rev. Lett.* 120.21 (2018), p. 211802. DOI: 10.1103/PhysRevLett.120.211802. arXiv: 1802.04329 [hep-ex].

[24] Morad Aaboud et al. "Search for the dimuon decay of the Higgs boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. Lett.* 119.5 (2017), p. 051802. DOI: 10.1103/PhysRevLett.119.051802. arXiv: 1705.04582 [hep-ex].

[25] Georges Aad et al. "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV". In: *JHEP* 08 (2016), p. 045. DOI: 10.1007/JHEP08(2016)045. arXiv: 1606.02266 [hep-ex].

[26] LHC Higgs Cross Section Working Group et al. "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties". In: *CERN-2013-004* (CERN, Geneva, 2013). DOI: 10.5170/CERN-2013-004. arXiv: 1307.1347 [hep-ph].

[27] Georges Aad et al. "Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector". In: *JHEP* 04 (2015), p. 117. DOI: 10.1007/JHEP04(2015)117. arXiv: 1501.04943 [hep-ex].

[28] Morad Aaboud et al. "Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector". In: *Phys. Lett.* B786 (2018), pp. 59–86. DOI: 10.1016/j.physletb.2018.09.013. arXiv: 1808.08238 [hep-ex].

[29] Albert M Sirunyan et al. "Observation of the Higgs boson decay to a pair of tau leptons". In: (2017). arXiv: 1708.00373 [hep-ex].

[30] A. M. Sirunyan et al. "Observation of Higgs boson decay to bottom quarks". In: *Phys. Rev. Lett.* 121.12 (2018), p. 121801. DOI: 10.1103/PhysRevLett.121.121801. arXiv: 1808.08242 [hep-ex].

[31] Albert M Sirunyan et al. "Combined measurements of Higgs boson couplings in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Submitted to: Eur. Phys. J.* (2018). arXiv: 1809.10733 [hep-ex].

[32] Georges Aad et al. "Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector". In: *Eur. Phys. J.* C75.10 (2015). [Erratum: Eur. Phys. J.C76,no.3,152(2016)], p. 476. DOI: 10.1140/epjc/s10052-015-3685-1,10.1140/epjc/s10052-016-3934-y. arXiv: 1506.05669 [hep-ex].

[33] A. D. Sakharov. "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe". In: *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967). [Usp. Fiz. Nauk161,no.5,61(1991)], pp. 32–35. DOI: 10.1070/PU1991v034n05ABEH002497.

[34] Stephen P. Martin. "A Supersymmetry primer". In: (1997). [Adv. Ser. Direct. High Energy Phys.18,1(1998)], pp. 1–98. DOI: 10.1142/9789812839657_0001,10.1142/9789814307505_0001. arXiv: hep-ph/9709356 [hep-ph].

[35] Abdelhak Djouadi. "The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model". In: *Phys. Rept.* 457 (2008), pp. 1–216. DOI: 10.1016/j.physrep.2007.10.004. arXiv: hep-ph/0503172 [hep-ph].

[36] Abdelhak Djouadi. "The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model". In: *Phys. Rept.* 459 (2008), pp. 1–241. DOI: 10.1016/j.physrep.2007.10.005. arXiv: hep-ph/0503173 [hep-ph].

[37] E. Accomando et al. "Workshop on CP Studies and Non-Standard Higgs Physics". In: (2006). DOI: 10.5170/CERN-2006-009. arXiv: hep-ph/0608079 [hep-ph].

[38] CMS Collaboration. "Measurement of the properties of a Higgs boson in the four-lepton final state". In: *Phys. Rev. D* 89 (2014), p. 092007. DOI: 10.1103/PhysRevD.89.092007. arXiv: 1312.5353 [hep-ex].

[39] Serguei Chatrchyan et al. "Measurement of the properties of a Higgs boson in the four-lepton final state". In: *Phys. Rev.* D89.9 (2014), p. 092007. DOI: 10.1103/PhysRevD.89.092007. arXiv: 1312.5353 [hep-ex].

[40] CMS Collaboration. "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at (7) and (8TeV)". In: *Eur. Phys. J. C* 75 (2015), p. 212. DOI: 10.1140/epjc/s10052-015-3351-7. arXiv: 1412.8662 [hep-ex].

[41] V. Khachatryan et al. "Combined search for anomalous pseudoscalar HVV couplings in VH(H $\to b\bar{b}$) production and H $\to$ VV decay". In: *Phys. Lett.* B759 (2016), pp. 672–696. DOI: 10.1016/j.physletb.2016.06.004. arXiv: 1602.04305 [hep-ex].

[42] Albert M Sirunyan et al. "Constraints on anomalous Higgs boson couplings using production and decay information in the four-lepton final state". In: (2017). arXiv: 1707.00541 [hep-ex].

[43] Vardan Khachatryan et al. "Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV". In: *Phys. Rev.* D92.1 (2015), p. 012004. DOI: 10.1103/PhysRevD.92.012004. arXiv: 1411.3441 [hep-ex].

[44] Vardan Khachatryan et al. "Limits on the Higgs boson lifetime and width from its decay to four charged leptons". In: *Phys. Rev.* D92.7 (2015), p. 072010. DOI: 10.1103/PhysRevD.92.072010. arXiv: 1507.06656 [hep-ex].

[45] Georges Aad et al. "Evidence for the spin-0 nature of the Higgs boson using ATLAS data". In: *Phys. Lett.* B726 (2013), pp. 120–144. DOI: 10.1016/j.physletb.2013.08.026. arXiv: 1307.1432 [hep-ex].

[46] ATLAS Collaboration. "Determination of spin and parity of the Higgs boson in the $WW^* \to e\nu\mu\nu$ decay channel with the ATLAS detector". In: *Eur. Phys. J. C* 75 (2015), p. 231. DOI: 10.1140/epjc/s10052-015-3436-3. arXiv: 1503.03643 [hep-ex].

[47] ATLAS Collaboration. "Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector". In: *Eur. Phys. J. C* 75 (2015), p. 476. DOI: 10.1140/epjc/s10052-015-3685-1. arXiv: 1506.05669 [hep-ex].

[48] M. Kramer et al. "Prospects of measuring the parity of Higgs particles". In: *Z. Phys.* C64 (1994), pp. 21–30. DOI: 10.1007/BF01557231. arXiv: hep-ph/9404280 [hep-ph].

[49] G. R. Bower et al. "Measuring the Higgs boson's parity using tau —¿ rho nu". In: *Phys. Lett.* B543 (2002), pp. 227–234. DOI: 10.1016/S0370-2693(02)02445-0. arXiv: hep-ph/0204292 [hep-ph].

[50] K. Desch et al. "Probing the CP nature of the Higgs boson at linear colliders with tau spin correlations: The Case of mixed scalar - pseudoscalar couplings". In: *Phys. Lett.* B579 (2004), pp. 157–164. DOI: 10.1016/j.physletb.2003.10.074. arXiv: hep-ph/0307331 [hep-ph].

[51] Andre Rouge. "CP violation in a light Higgs boson decay from tau-spin correlations at a linear collider". In: *Phys. Lett.* B619 (2005), pp. 43–49. DOI: 10.1016/j.physletb.2005.05.076. arXiv: hep-ex/0505014 [hep-ex].

[52] Stefan Berge, Werner Bernreuther, and Sebastian Kirchner. "Prospects of constraining the Higgs boson's CP nature in the tau decay channel at the LHC". In: *Phys. Rev.* D92 (2015), p. 096012. DOI: 10.1103/PhysRevD.92.096012. arXiv: 1510.03850 [hep-ph].

[53] J. D. Bjorken and Steven Weinberg. "A Mechanism for Nonconservation of Muon Number". In: *Phys. Rev. Lett.* 38 (1977), p. 622. DOI: 10.1103/PhysRevLett.38.622.

[54] Jong-Phil Lee and Kang Young Lee. "$B_s \to \mu\tau$ and $h \to \mu\tau$ decays in the general two Higgs doublet model". In: (2016). arXiv: 1612.04057 [hep-ph].

[55] J. Lorenzo Diaz-Cruz and J. J. Toscano. "Lepton flavor violating decays of Higgs bosons beyond the standard model". In: *Phys. Rev.* D62 (2000), p. 116005. DOI: 10.1103/PhysRevD.62.116005. arXiv: hep-ph/9910233 [hep-ph].

[56] Tao Han and Danny Marfatia. "h $\to$ mu tau at hadron colliders". In: *Phys. Rev. Lett.* 86 (2001), pp. 1442–1445. DOI: 10.1103/PhysRevLett.86.1442. arXiv: hep-ph/0008141 [hep-ph].

[57] Ernesto Arganda et al. "Lepton flavor violating Higgs boson decays from massive seesaw neutrinos". In: *Phys. Rev.* D71 (2005), p. 035011. DOI: 10.1103/PhysRevD.71.035011. arXiv: hep-ph/0407302 [hep-ph].

[58] Abdesslam Arhrib, Yifan Cheng, and Otto C. W. Kong. "Comprehensive analysis on lepton flavor violating Higgs boson to $\mu^{\mp}\tau^{\pm}$ decay in supersymmetry without R parity". In: *Phys. Rev.* D87.1 (2013), p. 015025. DOI: 10.1103/PhysRevD.87.015025. arXiv: 1210.8241 [hep-ph].

[59] M. Arana-Catania, E. Arganda, and M. J. Herrero. "Non-decoupling SUSY in LFV Higgs decays: a window to new physics at the LHC". In: *JHEP* 09 (2013). [Erratum: JHEP10,192(2015)], p. 160. DOI: 10.1007/JHEP10(2015)192, 10.1007/JHEP09(2013)160. arXiv: 1304.3371 [hep-ph].

[60] Kaustubh Agashe and Roberto Contino. "Composite Higgs-Mediated FCNC". In: *Phys. Rev.* D80 (2009), p. 075016. DOI: 10.1103/PhysRevD.80.075016. arXiv: 0906.1542 [hep-ph].

[61] Aleksandr Azatov, Manuel Toharia, and Lijun Zhu. "Higgs Mediated FCNC's in Warped Extra Dimensions". In: *Phys. Rev.* D80 (2009), p. 035016. DOI: 10.1103/PhysRevD.80.035016. arXiv: 0906.1990 [hep-ph].

[62] Gilad Perez and Lisa Randall. "Natural Neutrino Masses and Mixings from Warped Geometry". In: *JHEP* 01 (2009), p. 077. DOI: 10.1088/1126-6708/2009/01/077. arXiv: 0805.4652 [hep-ph].

[63] A. M. Baldini et al. "Search for the lepton flavour violating decay $\mu^+ \to e^+\gamma$ with the full dataset of the MEG experiment". In: *Eur. Phys. J.* C76.8 (2016), p. 434. DOI: 10.1140/epjc/s10052-016-4271-x. arXiv: 1605.05081 [hep-ex].

[64] K. Hayasaka et al. "Search for Lepton Flavor Violating Tau Decays into Three Leptons with 719 Million Produced Tau+Tau- Pairs". In: *Phys. Lett.* B687 (2010), pp. 139–143. DOI: 10.1016/j.physletb.2010.03.037. arXiv: 1001.3221 [hep-ex].

[65] Roni Harnik, Joachim Kopp, and Jure Zupan. "Flavor Violating Higgs Decays". In: *JHEP* 03 (2013), p. 026. DOI: 10.1007/JHEP03(2013)026. arXiv: 1209.1397 [hep-ph].

[66] Vardan Khachatryan et al. "Search for lepton flavour violating decays of the Higgs boson to $e\tau$ and $e\mu$ in proton–proton collisions at $\sqrt{s} = 8$ TeV". In: *Phys. Lett.* B763 (2016), pp. 472–500. DOI: 10.1016/j.physletb.2016.09.062. arXiv: 1607.03561 [hep-ex].

[67] U. Bellgardt et al. "Search for the Decay mu+ → e+ e+ e-". In: *Nucl. Phys.* B299 (1988), pp. 1–6. DOI: 10.1016/0550-3213(88)90462-2.

[68] Albert M Sirunyan et al. "Search for lepton flavour violating decays of the Higgs boson to $\mu\tau$ and eτ in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 06 (2018), p. 001. DOI: 10.1007/JHEP06(2018)001. arXiv: 1712.07173 [hep-ex].

[69] Bernard Aubert et al. "Searches for Lepton Flavor Violation in the Decays $\tau^\pm \to e^\pm\gamma$ and $\tau^\pm \to \mu^\pm\gamma$". In: *Phys. Rev. Lett.* 104 (2010), p. 021802. DOI: 10.1103/PhysRevLett.104.021802. arXiv: 0908.2381 [hep-ex].

[70] Lorenzo Calibbi and Giovanni Signorelli. "Charged Lepton Flavour Violation: An Experimental and Theoretical Introduction". In: *Riv. Nuovo Cim.* 41.2 (2018), p. 1. DOI: 10.1393/ncr/i2018-10144-0. arXiv: 1709.00294 [hep-ph].

[71] Georges Aad et al. "Search for lepton-flavour-violating $H \to \mu\tau$ decays of the Higgs boson with the ATLAS detector". In: *JHEP* 11 (2015), p. 211. DOI: 10.1007/JHEP11(2015)211. arXiv: 1508.03372 [hep-ex].

[72] Georges Aad et al. "Search for lepton-flavour-violating decays of the Higgs and Z bosons with the ATLAS detector". In: *Eur. Phys. J.* C77.2 (2017), p. 70. DOI: 10.1140/epjc/s10052-017-4624-0. arXiv: 1604.07730 [hep-ex].

[73] Vardan Khachatryan et al. "Search for Lepton-Flavour-Violating Decays of the Higgs Boson". In: *Phys. Lett.* B749 (2015), pp. 337–362. DOI: 10.1016/j.physletb.2015.07.053. arXiv: 1502.07400 [hep-ex].

[74] G L Bayatian et al. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. There is an error on cover due to a technical problem for some items. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/922757.

[75] A. Augusto Alves Jr. et al. "The LHCb Detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[76] K. Aamodt et al. "The ALICE experiment at the CERN LHC". In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[77] Christiane Lefèvre. "The CERN accelerator complex. Complexe des accelerateurs du CERN". Dec. 2008. URL: https://cds.cern.ch/record/1260465.

[78] *Luminosity Public Results Run2*. Feb. 12, 2019. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 (visited on 02/12/2019).

[79] Georges Aad et al. "Improved luminosity determination in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J.* C73.8 (2013), p. 2518. DOI: 10.1140/epjc/s10052-013-2518-3. arXiv: 1302.4393 [hep-ex].

[80] Morad Aaboud et al. "Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J.* C76.12 (2016), p. 653. DOI: 10.1140/epjc/s10052-016-4466-1. arXiv: 1608.03953 [hep-ex].

[81] Morad Aaboud et al. "Performance of the ATLAS Trigger System in 2015". In: *Eur. Phys. J.* C77.5 (2017), p. 317. DOI: 10.1140/epjc/s10052-017-4852-3. arXiv: 1611.09661 [hep-ex].

[82] *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*. Tech. rep. ATLAS-CONF-2016-024. Geneva: CERN, June 2016. URL: https://cds.cern.ch/record/2157687.

[83] *Egamma Trigger Public Results*. Feb. 12, 2019. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults (visited on 02/12/2019).

[84] Aranzazu Ruiz-Martinez et al. *Electron and photon trigger efficiency plots using the full 2016 dataset*. Tech. rep. ATL-COM-DAQ-2017-015. Geneva: CERN, Mar. 2017. URL: https://cds.cern.ch/record/2254973.

[85] Georges Aad et al. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV". In: *Eur. Phys. J.* C76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[86] Lidia Dell'Asta and Savanna Marie Shaw. *Muon trigger public plots with full 2016 dataset*. Tech. rep. ATL-COM-DAQ-2017-005. Geneva: CERN, Jan. 2017. URL: https://cds.cern.ch/record/2242824.

[87] Georges Aad et al. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J.* C77 (2017), p. 490. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex].

[88] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The Anti-k(t) jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[89] M. Aaboud et al. "Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev.* D96.7 (2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. arXiv: 1703.09665 [hep-ex].

[90] Morad Aaboud et al. "Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector". In: *Eur. Phys. J.* C77.9 (2017). [Erratum: Eur. Phys. J.C77,no.10,712(2017)], p. 580. DOI: 10.1140/epjc/s10052-017-5081-5,10.1140/epjc/s10052-017-5245-3. arXiv: 1705.02211 [hep-ex].

[91] Georges Aad et al. "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector". In: *Eur. Phys. J.* C76.11 (2016), p. 581. DOI: 10.1140/epjc/s10052-016-4395-z. arXiv: 1510.03823 [hep-ex].

[92] Morad Aaboud et al. "Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV". In: *JHEP* 08 (2018), p. 089. DOI: 10.1007/JHEP08(2018)089. arXiv: 1805.01845 [hep-ex].

[93] Morad Aaboud et al. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J.* C78.11 (2018), p. 903. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex].

[94] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[95] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[96] T Barillari et al. *Local Hadronic Calibration*. Tech. rep. ATL-LARG-PUB-2009-001-2. ATL-COM-LARG-2008-006. ATL-LARG-PUB-2009-001. Due to a report-number conflict with another document, the report-number ATL-LARG-PUB-2009-001-2 has been assigned. Geneva: CERN, June 2008. URL: https://cds.cern.ch/record/1112035.

[97] *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*. Tech. rep. ATL-PHYS-PUB-2015-045. Geneva: CERN, Nov. 2015. URL: https://cds.cern.ch/record/2064383.

[98] Georges Aad et al. "Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}$=8 TeV". In: *Eur. Phys. J.* C75.7 (2015), p. 303. DOI: 10.1140/epjc/s10052-015-3500-z. arXiv: 1412.7086 [hep-ex].

[99] Johann H. Kuhn and A. Santamaria. "Tau decays to pions". In: *Z. Phys.* C48 (1990), pp. 445–452. DOI: 10.1007/BF01572024.

[100] N. Davidson et al. "Universal Interface of TAUOLA Technical and Physics Documentation". In: *Comput. Phys. Commun.* 183 (2012), pp. 821–843. DOI: 10.1016/j.cpc.2011.12.009. arXiv: 1002.0543 [hep-ph].

[101] U. Baur. "Measuring the W boson mass at hadron colliders". In: *Electroweak precision data and the Higgs mass. Proceedings, Workshop, Zeuthen, Germany, February 28-March 1, 2003*. 2003, pp. 47–59. arXiv: hep-ph/0304266 [hep-ph]. URL: http://www-library.desy.de/preparch/desy/proc//proc03-01/5.ps.gz.

[102] Paolo Nason. "A new method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146.

[103] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with parton shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[104] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[105] E. Bagnaschi et al. "Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM". In: *JHEP* 02 (2012), p. 088. DOI: 10.1007/JHEP02(2012)088. arXiv: 1111.2854 [hep-ph].

[106] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "PYTHIA 6.4 Physics and Manual". In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph].

[107] T. Sjostrand, S. Mrenna and P. Skands. "A Brief Introduction to PYTHIA 8.1". In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[108] Marco Guzzi et al. "CT10 parton distributions and other developments in the global QCD analysis". In: (2011). arXiv: 1101.0561 [hep-ph].

[109] Peter Zeiler Skands. "Tuning Monte Carlo Generators: The Perugia Tunes". In: *Phys. Rev.* D82 (2010), p. 074018. DOI: 10.1103/PhysRevD.82.074018. arXiv: 1005.3457 [hep-ph].

[110] *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at* $\sqrt{s} = 13$ TeV. Tech. rep. ATLAS-CONF-2017-029. Geneva: CERN, May 2017. URL: http://cds.cern.ch/record/2261772.

[111] *The ATLAS Tau Trigger in Run 2*. Tech. rep. ATLAS-CONF-2017-061. Geneva: CERN, July 2017. URL: http://cds.cern.ch/record/2274201.

[112] Keith Hamilton et al. "NNLOPS simulation of Higgs boson production". In: *JHEP* 10 (2013), p. 222. DOI: 10.1007/JHEP10(2013)222. arXiv: 1309.0017 [hep-ph].

[113] Keith Hamilton, Paolo Nason, and Giulia Zanderighi. "Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator". In: *JHEP* 05 (2015), p. 140. DOI: 10.1007/JHEP05(2015)140.

[114] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[115] Charalampos Anastasiou et al. "Higgs Boson Gluon-Fusion Production in QCD at Three Loops". In: *Phys. Rev. Lett.* 114 (2015), p. 212001. DOI: 10.1103/PhysRevLett.114.212001. arXiv: 1503.06056 [hep-ph].

[116] Charalampos Anastasiou et al. "High precision determination of the gluon fusion Higgs boson cross-section at the LHC". In: *JHEP* 05 (2016), p. 058. DOI: 10.1007/JHEP05(2016)058. arXiv: 1602.00695 [hep-ph].

[117] Stefano Actis et al. "NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders". In: *Phys. Lett.* B670 (2008), pp. 12–17. DOI: 10.1016/j.physletb.2008.10.018. arXiv: 0809.1301 [hep-ph].

[118] Charalampos Anastasiou, Radja Boughezal, and Frank Petriello. "Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion". In: *JHEP* 04 (2009), p. 003. DOI: 10.1088/1126-6708/2009/04/003. arXiv: 0811.3458 [hep-ph].

[119] M. Ciccolini, Ansgar Denner, and S. Dittmaier. "Strong and electroweak corrections to the production of Higgs + 2jets via weak interactions at the LHC". In: *Phys. Rev. Lett.* 99 (2007), p. 161803. DOI: 10.1103/PhysRevLett.99.161803. arXiv: 0707.0381 [hep-ph].

[120] Mariano Ciccolini, Ansgar Denner, and Stefan Dittmaier. "Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC". In: *Phys. Rev.* D77 (2008), p. 013002. DOI: 10.1103/PhysRevD.77.013002. arXiv: 0710.4749 [hep-ph].

[121] Paolo Bolzoni et al. "Higgs production via vector-boson fusion at NNLO in QCD". In: *Phys. Rev. Lett.* 105 (2010), p. 011801. DOI: 10.1103/PhysRevLett.105.011801. arXiv: 1003.4451 [hep-ph].

[122] Oliver Brein, Abdelhak Djouadi, and Robert Harlander. "NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders". In: *Phys. Lett.* B579 (2004), pp. 149–156. DOI: 10.1016/j.physletb.2003.10.112. arXiv: hep-ph/0307206 [hep-ph].

[123] Lukas Altenkamp et al. "Gluon-induced Higgs-strahlung at next-to-leading order QCD". In: *JHEP* 02 (2013), p. 078. DOI: 10.1007/JHEP02(2013)078. arXiv: 1211.5015 [hep-ph].

[124] Ansgar Denner et al. "Electroweak corrections to Higgs-strahlung off W/Z bosons at the Tevatron and the LHC with HAWK". In: *JHEP* 03 (2012), p. 075. DOI: 10.1007/JHEP03(2012)075. arXiv: 1112.5142 [hep-ph].

[125] ATLAS Collaboration. "Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *JHEP* 09 (2014), p. 145. DOI: 10.1007/JHEP09(2014)145. arXiv: 1406.3660 [hep-ex].

[126] J. Pumplin et al. "New generation of parton distributions with uncertainties from global QCD analysis". In: *JHEP* 07 (2002), p. 012. DOI: 10.1088/1126-6708/2002/07/012. arXiv: hep-ph/0201195.

[127] R. Ball et al. "Parton distributions for the LHC Run II". In: *JHEP* 1504 (2014), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[128] Tanju Gleisberg and Stefan Hoche. "Comix, a new matrix element generator". In: *JHEP* 12 (2008), p. 039. DOI: 10.1088/1126-6708/2008/12/039. arXiv: 0808.3674 [hep-ph].

[129] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini. "Scattering Amplitudes with Open Loops". In: *Phys. Rev. Lett.* 108 (2012), p. 111601. DOI: 10.1103/PhysRevLett.108.111601. arXiv: 1111.5206 [hep-ph].

[130] Kirill Melnikov and Frank Petriello. "Electroweak gauge boson production at hadron colliders through O(alpha(s)**2)". In: *Phys. Rev.* D74 (2006), p. 114017. DOI: 10.1103/PhysRevD.74.114017. arXiv: hep-ph/0609070 [hep-ph].

[131] Charalampos Anastasiou et al. "High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO". In: *Phys. Rev.* D69 (2004), p. 094008. DOI: 10.1103/PhysRevD.69.094008. arXiv: hep-ph/0312266 [hep-ph].

[132] Simone Alioli et al. "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions". In: *JHEP* 09 (2009). [Erratum: JHEP02,011(2010)], p. 111. DOI: 10.1007/JHEP02(2010)011,10.1088/1126-6708/2009/09/111. arXiv: 0907.4076 [hep-ph].

[133] Emanuele Re. "Single-top Wt-channel production matched with parton showers using the POWHEG method". In: *Eur. Phys. J.* C71 (2011), p. 1547. DOI: 10.1140/epjc/s10052-011-1547-z. arXiv: 1009.2450 [hep-ph].

[134] J. Pumplin et al. "New generation of parton distributions with uncertainties from global QCD analysis". In: *JHEP* 07 (2002), p. 012. DOI: 10.1088/1126-6708/2002/07/012. arXiv: hep-ph/0201195 [hep-ph].

[135] Michal Czakon and Alexander Mitov. "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders". In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: 10.1016/j.cpc.2014.06.021. arXiv: 1112.5675 [hep-ph].

[136] Georges Aad et al. "Modelling $Z \to \tau\tau$ processes in ATLAS with τ-embedded $Z \to \mu\mu$ data". In: *JINST* 10.09 (2015), P09018. DOI: 10.1088/1748-0221/10/09/P09018, 10.1088/1748-0221/2015/9/P09018. arXiv: 1506.05623 [hep-ex].

[137] A. Elagin et al. "A New Mass Reconstruction Technique for Resonances Decaying to di-tau". In: *Nucl. Instrum. Meth.* A654 (2011), pp. 481–489. DOI: 10.1016/j.nima.2011.07.009. arXiv: 1012.4686 [hep-ex].

[138] *Electron identification measurements in ATLAS using $\sqrt{s}$ = 13 TeV data with 50 ns bunch spacing*. Tech. rep. ATL-PHYS-PUB-2015-041. Geneva: CERN, Sept. 2015. URL: https://cds.cern.ch/record/2048202.

[139] Morad Aaboud et al. "Electron and photon energy calibration with the ATLAS detector using 2015-2016 LHC proton-proton collision data". In: (2018). arXiv: 1812.03848 [hep-ex].

[140] *E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC p-p collisions*. Tech. rep. ATLAS-CONF-2018-023. Geneva: CERN, June 2018. URL: http://cds.cern.ch/record/2625233.

[141] D. de Florian et al. "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector". In: (2016). DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph].

[142] Jon Butterworth et al. "PDF4LHC recommendations for LHC Run II". In: *J. Phys.* G43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph].

[143] Leif Lonnblad and Stefan Prestel. "Matching Tree-Level Matrix Elements with Interleaved Showers". In: *JHEP* 03 (2012), p. 019. DOI: 10.1007/JHEP03(2012)019. arXiv: 1109.4829 [hep-ph].

[144] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. San Francisco, California, USA: ACM, 2016, pp. 785–794. ISBN: 978-1-4503-4232-2. DOI: 10.1145/2939672.2939785. URL: http://doi.acm.org/10.1145/2939672.2939785.

[145] Ruben Martinez-Cantin. "BayesOpt: A Bayesian Optimization Library for Nonlinear Optimization, Experimental Design and Bandits". In: *Journal of Machine Learning Research* 15 (2014), pp. 3915–3919. URL: http://jmlr.org/papers/v15/martinezcantin14a.html.

[146] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference and prediction*. 2nd ed. Springer, 2009. URL: http://www-stat.stanford.edu/~tibs/ElemStatLearn/.

[147] R. Jozefowicz, E. Richter-Was, and Z. Was. "Potential for optimizing the Higgs boson CP measurement in H $\to \tau\tau$ decays at the LHC including machine learning techniques". In: *Phys. Rev.* D94.9 (2016), p. 093001. DOI: 10.1103/PhysRevD.94.093001. arXiv: 1608.02609 [hep-ph].

[148] Stefan Berge, Werner Bernreuther, and Sebastian Kirchner. "Determination of the Higgs CP-mixing angle in the tau decay channels at the LHC including the Drell–Yan background". In: *Eur. Phys. J.* C74.11 (2014), p. 3164. DOI: 10.1140/epjc/s10052-014-3164-0. arXiv: 1408.0798 [hep-ph].

[149] S. Berge et al. "How to pin down the CP quantum numbers of a Higgs boson in its tau decays at the LHC". In: *Phys. Rev.* D84 (2011), p. 116003. DOI: 10.1103/PhysRevD.84.116003. arXiv: 1108.0670 [hep-ph].

[150] Stefan Berge and Werner Bernreuther. "Determining the CP parity of Higgs bosons at the LHC in the tau to 1-prong decay channels". In: *Phys. Lett.* B671 (2009), pp. 470–476. DOI: 10.1016/j.physletb.2008.12.065. arXiv: 0812.1910 [hep-ph].

[151] Stefan Berge, Werner Bernreuther, and Sebastian Kirchner. "Determination of the Higgs CP-mixing angle in the tau decay channels". In: *Nucl. Part. Phys. Proc.* 273-275 (2016), pp. 841–845. DOI: 10.1016/j.nuclphysbps.2015.09.129. arXiv: 1410.6362 [hep-ph].

[152] Stanislaw Jadach, J H Kühn, and Zbigniew Was. "TAUOLA: a library of Monte Carlo programs to simulate decays of polarized τ leptons". In: *Comput. Phys. Commun.* 64.CERN-TH-5856-90 (Sept. 1990), 275–299. 38 p. URL: https://cds.cern.ch/record/212328.

[153] Simonetta Gentile and Martin Pohl. "Physics of tau leptons". In: *Physics Reports* 274.5 (1996), pp. 287–374. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(96)00002-6. URL: http://www.sciencedirect.com/science/article/pii/0370157396000026.

[154] T. Przedzinski, E. Richter-Was, and Z. Was. "TauSpinner: a tool for simulating CP effects in $H \to \tau\tau$ decays at LHC". In: *Eur. Phys. J.* C74.11 (2014), p. 3177. DOI: 10.1140/epjc/s10052-014-3177-8. arXiv: 1406.1647 [hep-ph].

[155] Francois Chollet et al. *Keras*. https://github.com/fchollet/keras. 2015.

[156] Martin Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: http://tensorflow.org/.

[157] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning". In: *Nature* 521 (May 2015), URL: https://doi.org/10.1038/nature14539.

[158] D. P. Kingma and J. Ba. "Adam: A Method for Stochastic Optimization". In: *ArXiv e-prints* (Dec. 2014). arXiv: 1412.6980.

[159] Nitish Srivastava et al. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting". In: *Journal of Machine Learning Research* 15 (2014), pp. 1929–1958. URL: http://jmlr.org/papers/v15/srivastava14a.html.

[160] G. Aad et al. "The ATLAS Inner Detector commissioning and calibration". In: *Eur. Phys. J.* C70 (2010), pp. 787–821. DOI: 10.1140/epjc/s10052-010-1366-7. arXiv: 1004.5293 [physics.ins-det].

[161] Georges Aad et al. "Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment". In: *Eur. Phys. J.* C76.5 (2016), p. 295. DOI: 10.1140/epjc/s10052-016-4110-0. arXiv: 1512.05955 [hep-ex].

[162] B. Efron. *The Jackknife, the Bootstrap and Other Resampling Plans*. Society for Industrial and Applied Mathematics, 1982. DOI: 10.1137/1.9781611970319. eprint: http://epubs.siam.org/doi/pdf/10.1137/1.9781611970319. URL: http://epubs.siam.org/doi/abs/10.1137/1.9781611970319.

[163] D.M. Asner et al. "Hadronic structure in the decay $\tau^- \to \nu_\tau \pi^- \pi^0 \pi^0$ and the sign of the tau-neutrino helicity". In: *Phys.Rev.* D61 (2000), p. 012002. DOI: 10.1103/PhysRevD.61.012002. arXiv: hep-ex/9902022 [hep-ex].

[164] M. Chrzaszcz et al. "TAUOLA of tau lepton decays– framework for hadronic currens, matrix elements and anomalous decays". In: (2016). arXiv: 1609.04617 [hep-ph].

[165] J. W. Hinson. "Axial vector and pseudoscalar hadronic structure in $\tau \to \pi^- \pi^- \pi^+ \nu_\tau$ decays with implications on light quark masses". PhD thesis. Purdue Universty, 2001.

[166] Ian M. Nugent. *Precision Measurements of Tau Lepton Decays*. SLAC-R-936, Ph.D. Thesis.

[167] D. Gomez Dumm et al. "$\tau \to \pi\pi\pi\nu_\tau$ decays and the a(1)(1260) off-shell width revisited". In: *Phys. Lett.* B685 (2010), pp. 158–164. DOI: 10.1016/j.physletb.2010.01.059. arXiv: 0911.4436 [hep-ph].

[168] Florian Beaudette. "The CMS Particle Flow Algorithm". In: *Proceedings, International Conference on Calorimetry for the High Energy Frontier (CHEF 2013): Paris, France, April 22-25, 2013*. 2013, pp. 295–304. arXiv: 1401.8155 [hep-ex].

Appendices

A

Tau Decay Mode Identification

Thus far, the τ_{had} reconstruction has treated candidates as a jet-like object rather than being concerned with individual hadron responses. This can be considered a "baseline" reconstruction algorithm. A reconstruction method was developed in order to exploit all information from both tracking and calorimetry components in order to reconstruct decay products. The development of this new reconstruction is crucial in $H \to \tau\tau$ CP measurements, discussed later in chapter 7. The measurement of the CP state of the Higgs boson decaying to a pair of τ_{had} requires knowledge of the kinematic distributions of individual hadrons rather than of the τ lepton itself. In addition the measurment method varies depending on the number of charged and neutral pions in the final state.

The idea of the reconstruction of the decay products relies on a particle flow approach [168], integrating information from the tracking systems as well as calorimeter clusters, in contrast to the calorimeter only approach used to reconstruct the jet-like objects detailed in previous sections. The tracking information allows for identification of charged hadrons which can be matched to calorimeter deposits in the hadronic calorimeter. The remaining clusters are then used as inputs for neutral pion reconstruction. Using this information a more accurate and detailed reconstuction is realised.

A.1 Neutral Pion Identification

Neutral pions predominantly decay via $\pi^0 \to \gamma\gamma$, resulting in deposits in the EM calorimeter. The energy deposited in the EM calorimeter which is associated with the charged hadron is obtained as the difference between the momentum measured by the tracking system and the associated deposits in the HAD calorimeter. The remaining EM clusters after subtracting the EM deposits matched with the charged hadron are considered for identification of neutral pions.

The most difficult task in identifying decay modes is correctly counting the number of neutral pions. Due to the nature of the LHC collisions, many other background sources of EM deposits exist such as pileup. In order to most effectively identify the neutral pions, a BDT is utilised.

The algorithm is trained with simulated $Z \rightarrow \tau\tau$ events, where the τ_{had} are only 1p, to discriminate decays which do and do not produce neutral pions. Cluster topology information is used as input variables. The performance of the identification algorithm is presented in A.1 and A.2.

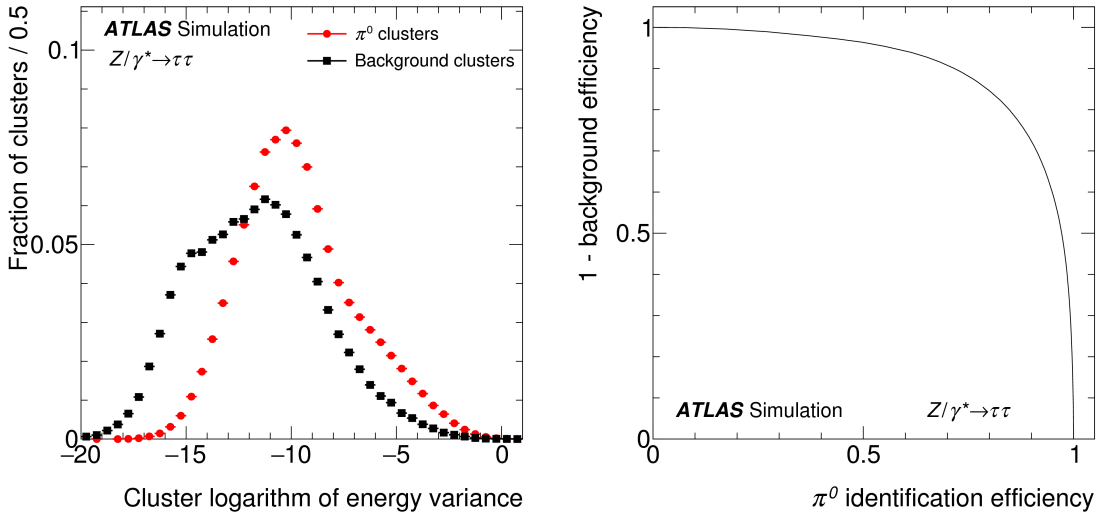


Figure A.1: The classifier output for π^0 identification BDT (left) and the receiver operating characteristic curve (right) [161].

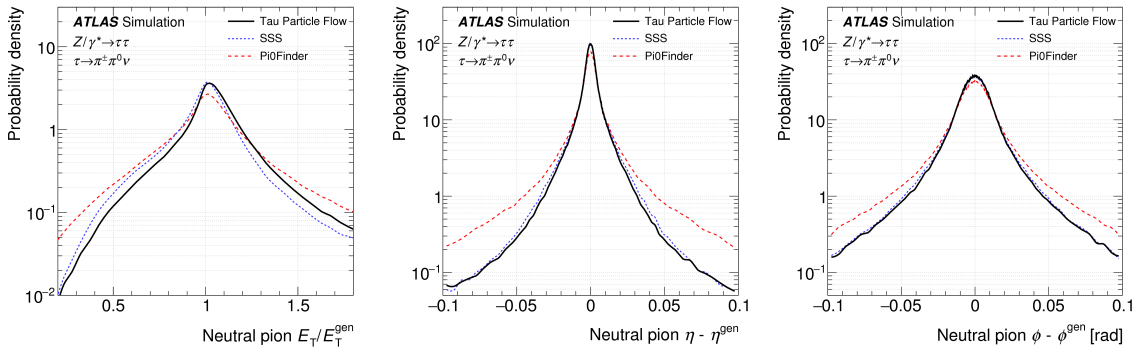


Figure A.2: The relative or fractional resolution of the reconstructed π^0 candidates in transverse energy (left), η (centre) and ϕ (right) [97]. The default algorithm is that of the Tau Particle Flow, which outperforms other algorithms outlined in [161].

A.2 Decay Mode Classification

With an algorithm for identifying pions, a classification scheme can be created to distinguish τ_{had} decay modes. The target is to identify five broad categories based on the number of

charged and neutral pions. The modes are 1p0n, 1p1n, 1pXn, 3p0n and 3pXn where p is the number of charged final state particles and n is the number of neutrals.[1] As the difficult task is to identify the number of neutral pions, three BDTs are trained to separate 1p0n vs 1p1n, 1p1n vs 1pXn and 3p0n vs 3pXn. The performance of the BDTs is summarised in figure A.3.

The four momentum can now be also be reconstructed for the total τ_{had}. Firstly, the four momenta of the identified constituents are summed. The particle flow approach then combines both the constituent and baseline reconstructed energies together. Each of the energies are weighted by the inverse square of the core resolution to provide the best resolution regardless of p_T. This weighting allows the particle flow approach to accurately model high energy taus where the baseline approach exceeds the accuracy of the constituent based approach. The resolution of the four momenta are depicted in A.4. Overall the reconstruction of the substructure allows for not only a reconstruction of the decay mode but, through reconstruction of decay products, an improved resolution.

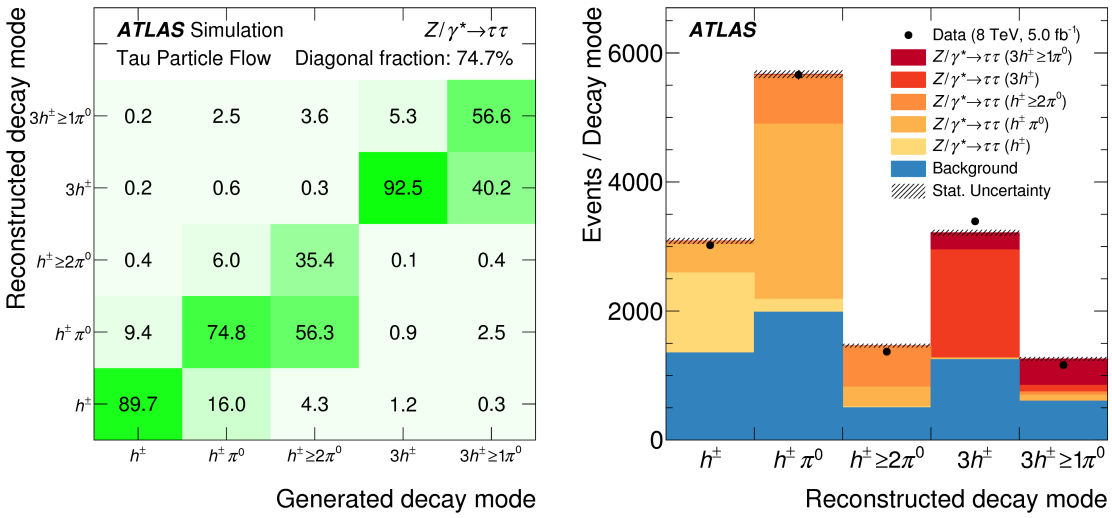


Figure A.3: The reconstruction efficies for each of the τ identification BDTs (left) and the receiver operating characteristic curve (right). Below shows the composition of reconstructed decay modes with a comparison with data [161].

[1] X represents > 1 or > 0 for 1p and 3p respectively.

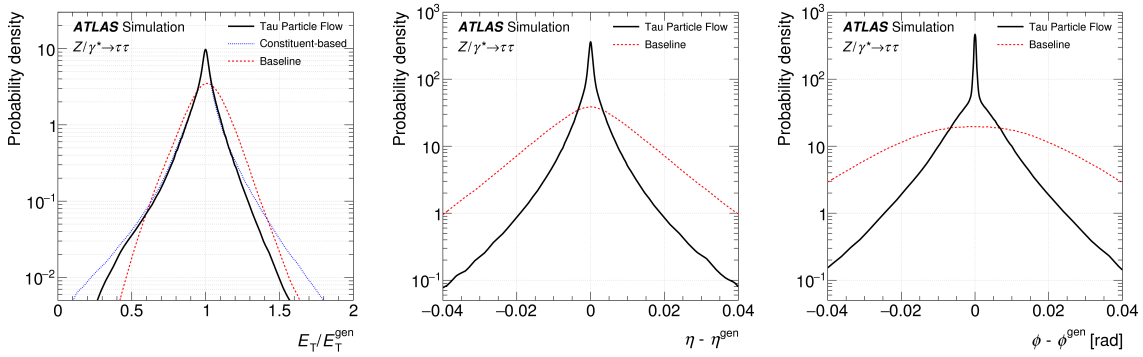


Figure A.4: The relative or fractional resolution of the reconstructed τ_{had} candidates in transverse energy (left), η (centre) and ϕ (right) [97]. The resolution with respect to the baseline reconstruction and the particle flow approach.

B

LFV MVA performance

The Bayesian optimisation scans are presented in figures B.1 and B.2.

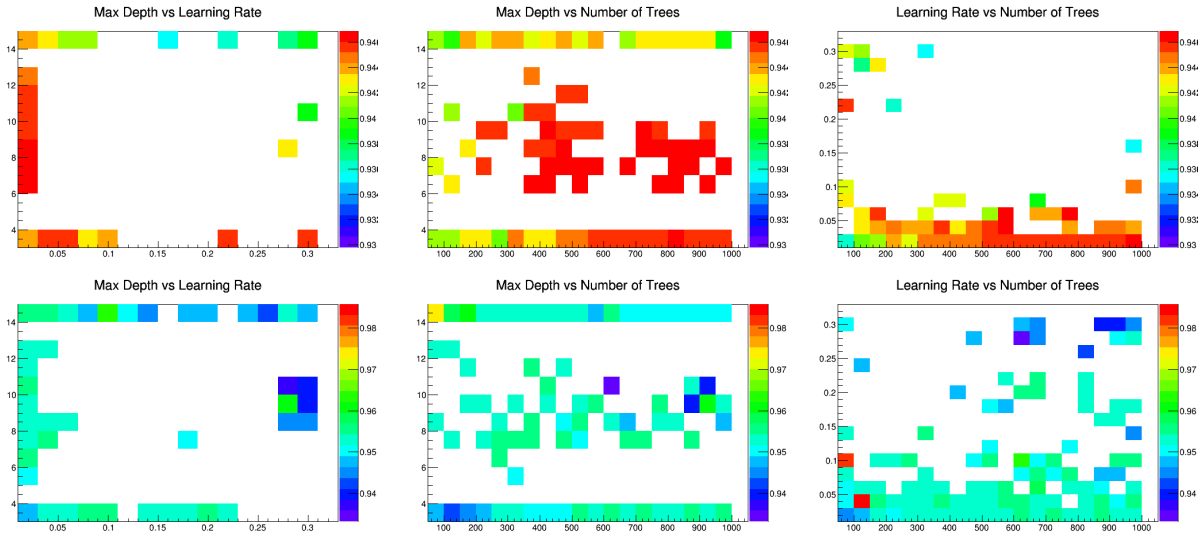


Figure B.1: 2D slices of the hyperparameters scanned over to train the BDTs. This scan is done via the Bayesian Optimisation method for the $e\tau_{had}$ channel

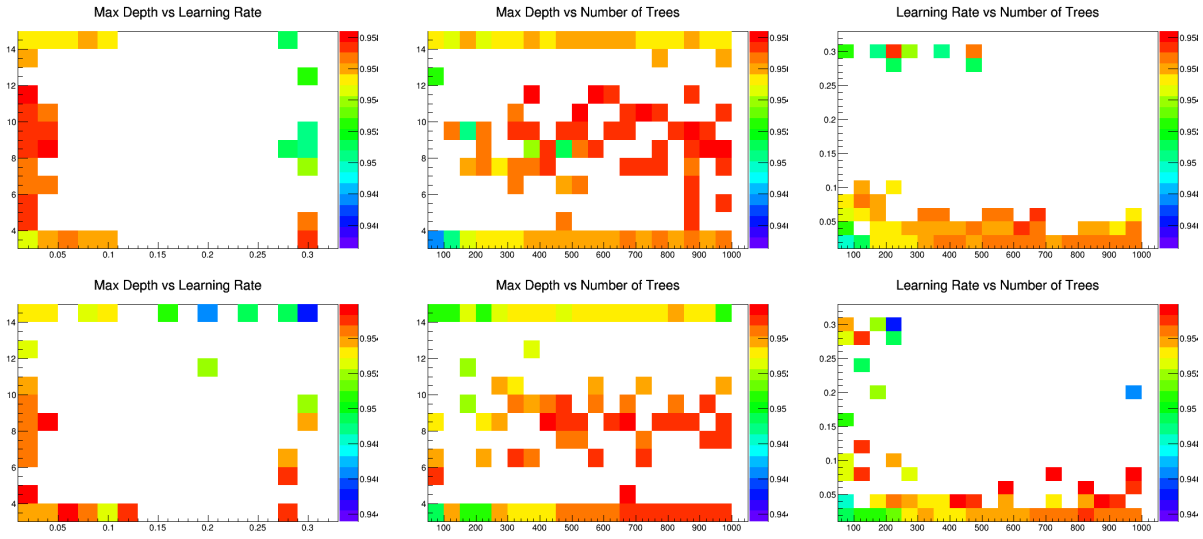


Figure B.2: AUC scores calculated on the validation dataset in 2D slices of the hyperparameters scanned over to train the BDTs. This scan is done via the Bayesian Optimisation method for the $e\tau_{had}$ channel